NATURAL LANGUAGE LEARNING BY COMPUTER

BY

LAURENT SIKLÓSSY

SUBMITTED TO THE CARNEGIE-MELLON UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

PITTSBURGH, PENNSYLVANIA

1968

THIS WORK WAS SUPPORTED PARTIALLY BY THE DANFORTH FOUNDATION AND PARTIALLY BY THE ADVANCED RESEARCH PROJECTS AGENCY OF THE OFFICE OF THE SECRETARY OF DEFENSE (SD-146) AND IS MONITORED BY THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH.

II

ABSTRACT

LEARNING A NATURAL LANGUAGE IS TAKEN AS AN IMPROVEMENT IN A SYSTEM'S ABILITY TO EXPRESS SITUATIONS IN A NATURAL LANGUAGE.

THIS DISSERTATION DESCRIBES A COMPUTER PROGRAM, CALLED ZBIE, WRITTEN IN IPL-V, WHICH ACCEPTS THE DESCRIPTION OF SITUATIONS IN A UNIFORM, STRUCTURED FUNCTIONAL LANGUAGE AND TRIES TO EXPRESS THESE SITUATIONS IN A NATURAL LANGUAGE. EXAMPLES ARE GIVEN FOR GERMAN AND, MOSTLY, RUSSIAN.

AT RUN-TIME, ZBIE BUILDS SIMPLE MEMORY STRUCTURES. PATTERNS AND SETS ARE BUILT ON THE FUNCTIONAL LANGUAGE. THE TRANSLATION RULES OF THE PATTERNS AND AN IN-CONTEXT VOCABULARY PROVIDE THE TRANSITION TO THE NATURAL LANGUAGE. ZBIE IS A CAUTIOUS LEARNER, AND AVOIDS ERRORS BY SEVERAL MECHANISMS. ZBIE IS CAPABLE OF SOME EVOLUTIONARY LEARNING.

ACKNOWLEDGMENTS

AMONG THE MANY STUDENTS WHO HAVE HELPED ME, I WOULD LIKE TO THANK PARTICULARLY G. BERGLASS, R. FIKES, P. FREEMAN, R. GROVE AND THE LOCAL IPL-V EXPERTS, R. BUSHYAGER AND T. CUNNINGHAM.

MY WIFE HAS SPENT MANY HOURS PROOFREADING THE THESIS AND HAS GIVEN ME MUCH MORAL SUPPORT.

ABOVE ALL I AM INDEBTED TO PROFESSOR HERBERT A. SIMON FOR HIS GUIDANCE AT ALL STAGES OF THIS WORK.

TABLE OF CONTENTS

CHAPTER I.

INTRODUCTION.

WORKERS IN ARTIFICIAL INTELLIGENCE HAVE, AS ONE OF THEIR GOALS, THE WRITING OF SOPHISTICATED COMPUTER PROGRAMS WHICH WILL PERFORM 'INTERESTING' AND 'DIFFICULT' TASKS. PROGRAMS CAN IMPROVE THEIR SOPHISTICATION BY LEARNING, AND LEARNING IS, INDEED, A CENTRAL PROBLEM OF ARTIFICIAL INTELLIGENCE.

ONE OF THE FIRST LEARNING TASKS THAT HUMAN BEINGS PERFORM IS ACQUIRING A NATURAL LANGUAGE (ABBREVIATED NL). THROUGHOUT HISTORY MEN HAVE USED NL'S FOR COMMUNICATING AMONG THEMSELVES AND INVESTIGATING AND INTERACTING WITH THE WORLD. FOR THE PAST DECADE, NATURAL LANGUAGE COMMUNICATION OF HUMANS WITH COMPUTERS HAS BEEN AN ACTIVE AREA OF INTEREST IN ARTIFICIAL INTELLIGENCE.

INTERESTS IN THE FIELDS OF LEARNING AND NATURAL LANGUAGE ARE COMBINED HERE IN A PROGRAM CALLED ZBIE THAT ATTEMPTS TO LEARN NATURAL LANGUAGES AT AN ELEMENTARY LEVEL. THE TASK IS CONSIDERED WORTHY OF INVESTIGATION IN ITS OWN RIGHT; THE PROGRAM DOES NOT TRY TO SIMULATE THE LEARNING BEHAVIOR OF HUMAN BEINGS.

NATURAL LANGUAGE LEARNING PROGRAMS HAVE BEEN FEW. TO THE BEST OF THE AUTHOR'S KNOWLEDGE, CORROBORATED BY A REVIEW OF

RECENT WORK IN ARTIFICIAL INTELLIGENCE (SOLOMONOFF 1966), ONLY ONE WORK CAN QUALIFY : UHR 1964. BY A PROCESS OF STRING MATCHING AND STATISTICAL LEARNING, UHR'S PROGRAMS ATTEMPT TO TRANSLATE STRINGS FROM ONE NL (NL1) INTO STRINGS OF ANOTHER NL (NL2). THE PROGRAMS ARE INSUFFICIENTLY DOCUMENTED TO EXPLAIN THEIR STRUCTURE IN DETAIL, BUT FROM THE OJTPUTS EXHIBITED, SEVERAL LIMITS APPEAR: THE IDIOSYNCRASIES OF NL1 GIVE DIFFICULTIES TO THE PROGRAMS AND THE LEARNING PROCESS SEEMS TO CYCLE.

A POSSIBLE CAUSE FOR THE LACK OF INTEREST IN NL LEARNING PROGRAMS IS THE FEELING, AMONG MANY LINGUISTS, THAT THE LANGUAGE LEARNING TASK IS EXTREMELY ARDUOUS. TWO OF THE FOREMOST SCIENTISTS IN THE FIELD OF MODERN LINGUISTICS STATED (CHOMSKY AND MILLER, 1963):

> TO IMAGINE THAT AN ADEQUATE GRAMMAR COULD BE SELECTED FROM THE INFINITUDE OF CONCEIVABLE ALTERNATIVES BY SOME PROCESS OF PURE INDUCTION ON A FINITE CORPUS OF UTTERANCES IS TO MISJUDGE COMPLETELY THE MAGNITUDE OF THE PROBLEM.

TWO RELATED AREAS HAVE RECEIVED MUCH MORE ATTENTION: THE INDUCTION OF GRAMMARS OF ABSTRACT LANGUAGES AND THE EXTRAPOLATION OF SEQUENCES. SOLOMONOFF (1958) OFFERED A SKETCH FOR THE MECHANIZATION OF LINGUISTIC LEARNING, WHICH DOES NOT APPEAR TO HAVE BEEN PROGRAMMED, AND, LATER (1964), HIS FORMAL THEORY OF INDUCTIVE INFERENCE PRESENTS VARIOUS MODELS FOR EXTRAPOLATING A LONG SEQUENCE OF SYMBOLS CONTAINING ALL DATA TO BE USED IN THE

PREDICTION. A PROBABILISTIC APPROACH IS USED. SIMON AND KOTOVSKY (1963) DESCRIBE PROGRAMS THAT INDUCT RULES TO EXPLAIN THE FORMATION OF GIVEN SERIAL ALPHABETIC PATTERNS. THE RULES CAN THEN BE APPLIED TO CONTINUE THE PATTERN. PIVAR AND FINKELSTEIN (1964) REPORTED ON A SIMILAR, PERFORMANCE-ORIENTED PROGRAM WHICH ALSO HANDLES INTEGER STRINGS. (FOR A REVIEW OF RECENT RESEARCH, SEE SOLOMONOFF, 1966). IN EITHER AREA, THE STRINGS CONSIDERED ARE 'EMPTY'; THEY LACK THE CONTENT THAT NATURAL LANGUAGES POSSESS AS WAYS OF EXPRESSING SITUATIONS OF OUR HUMAN ENVIRONMENT.

BEFORE TRYING TO DEFINE THE LEARNING TASK, LET US CONSIDER THE TECHNIQUE FOR TEACHING LANGUAGES (TO HUMANS) USED BY I. A. RICHARDS AND HIS CO-WORKERS (RICHARDS, 1961). IN THE LANGUAGE - THROUGH - PICTURES SERIES, PICTURES ARE ASSOCIATED WITH SENTENCES IN AN NL TO BE LEARNT. THE PICTURES ARE TO ACT AS A GENERAL REPRESENTATION FOR ALL HUMAN BEINGS ('ENGLISH THROUGH PICTURES, BOOK I' IS PREFACED IN 41 LANGUAGES). THE STUDENT IS SUPPOSED TO USE THE PICTURES AS CLUES TO THE MEANING OF THE SENTENCES AND, BY SUCCESSIVE COMPARISONS OF THE SENTENCES, TO INFER THE VOCABULARY AND GRAMMAR OF THE NL STUDIED.

THE STUDENT'S OWN MAIN OR MOTHER TONGUE IS BYPASSED, THEREBY AVOIDING PROBLEMS OF TRANSLATION FROM ONE TONGUE INTO ANOTHER; INSTEAD THE STUDENT LEARNS TO TRANSLATE SITUATIONS DIRECTLY FROM 'REALITY' INTO A NEW NL.

(AS AN ASIDE, THE AUTHOR MAY ADD THAT HE TRIED TO LEARN

HEBREW, ABSOLUTELY UNKNOWN BEFOREHAND, FROM 'HEBREW THROUGH PICTURES'. HE HAD THE ADVANTAGE OF HAVING HAD READ PREVIOUSLY SEVERAL OTHER 'LANGUAGE THROUGH PICTURES' BOOKS IN KNOWN LANGUAGES; NEVERTHELESS HE HAD GREAT DIFFICULTIES IN TRYING TO DETERMINE THE MEANINGS OF THE PICTURES OR THE CLUES TO BE DERIVED FROM THEM, AND HE FINALLY ABANDONED THE ENDEAVOUR. SEVERAL OTHER PERSONS REPORTED IDENTICAL DIFFICULTIES.)

THE PHILOSOPHIES BEHIND ZBIE AND I. A. RICHARDS'S BOOKLETS ARE SIMILAR. ZBIE USES A FUNCTIONAL LANGUAGE (ABBREVIATED FL) TO REPRESENT SITUATIONS; FL HAS THE SAME FUNCTION IN ZBIE AS THE PICTURES IN RICHARDS. BY SUCCESSIVE COMPARISONS OF SITUATIONS, AS REPRESENTED IN FL AND AS EXPRESSED IN AN NL, ZBIE TRIES TO LEARN HOW TO EXPRESS OTHER SITUATIONS REPRESENTED IN FL AND, FAILING THAT, TO USE ITS PREVIOUS KNOWLEDGE TO TRY TO LEARN HOW TO EXPRESS THE OTHER SITUATIONS. THE LEARNING SEQUENCE USED IS TAKEN FROM 'RUSSIAN THROUGH PICTURES' WITH SLIGHT MODIFICATIONS.

CHAPTER II IS DIVIDED INTO THREE PARTS.

IN PART A, WE DESCRIBE FL BRIEFLY.

IN PART B, WE DESCRIBE THE INTERNAL REPRESENTATIONS USED BY ZBIE: PATTERNS, SETS, TRANSLATION RULES AND IN-CONTEXT VOCABULARY.

IN PART C, WE DESCRIBE THE ORGANIZATION OF ZBIE AND THE MAIN PROCESSOR ROUTINES.

SINCE CHAPTER II IS RATHER DETAILED, THE READER MAY WELL WANT TO COME BACK TO IT AFTER READING THE FOLLOWING CHAPTER.

CHAPTER III COMMENTS ON ZBIE'S LEARNING OF RUSSIAN.

CHAPTER IV COMPARES ZBIE WITH UHR'S PROGRAMS AND DISCUSSES SOME OF ZBIE'S INADEQUACIES.

APPENDIX A SHOWS THE CODE USED TO TRANSLATE THE RUSSIAN CYRILLIC ALPHABET INTO LATIN ALPHANUMERICS.

APPENDIX B GIVES A SIMPLE EXAMPLE, IN GERMAN, OF ZBIE'S 'EVOLUTIONARY LEARNING' CAPABILITIES.

ZBIE IS CODED IN IPL-V (NEWELL, 1964) AND HAS BEEN RUN ON THE CARNEGIE-MELLON UNIVERSITY CDC G-21 COMPUTER. SINCE IPL-V CODE IS TYPICALLY UNREADABLE, THE PROGRAM IS NOT ENCLOSED, BUT IT IS DESCRIBED SEMI-FORMALLY IN CHAPTER II, PART C.

CHAPTER II.

A. THE FUNCTIONAL LANGUAGE, FL.

THE PURPOSE OF FL IS TO REPRESENT SITUATIONS IN A FASHION SOMEWHAT SIMILAR TO THE PICTURES (AND PICTURE LANGUAGE) USED IN THE LANGUAGE - THROUGH - PICTURES SERIES. THE MAIN SPIRIT BEHIND FL MAY BE SUMMARIZED THUS : 'SIMILAR SITUATIONS SHOULD HAVE SIMILAR REPRESENTATIONS IN FL', WHERE AN INTUITIVE FEELING FOR SIMILARITY IS USED. FOR EXAMPLE THE SENTENCES :

THIS IS A HAT.

THIS IS HIS HAT. (REFERRING TO A BOY)

THIS IS THE BOY'S HAT.

SHOULD HAVE SIMILAR REPRESENTATIONS IN FL. TO AVOID IDIOSYNCRASIES OF NL'S, FL IS NOT INFLECTED AND OMITS ARTICLES TO IMPROVE ITS DESCRIPTIVE POWER WE HAVE ADDED SOME SEMANTICS FOR INSTANCE THE REFERENT OF PRONOUNS IS SPECIFICALLY MENTIONED: HE=(MAN) OR YOU=(SPOKEN BOY) IF THE PERSON SPOKEN TO IS A BOY (REMEMBER THE PICTURES).

FL IS NOT UNLIKE THE LANGUAGE DESCRIBED BY REICHENBACH (1947) FOR HIS ANALYSIS OF ENGLISH. INSTEAD OF THE USUAL FUNCTIONAL NOTATION $F(X1, X2, \ldots, XN)$ WE USE A LISP-LIKE NOTATION

(MCCARTHY 1963), (F X1 X2 ... XN) AND ALSO USE DESCRIPTION LISTS, ENCLOSED IN SQUARE BRACKETS [AND].

VERBS AND FUNCTION WORDS ARE TREATED AS N-PLACE FUNCTIONS. A FEW EXAMPLES SHOULD MAKE SOME OF THE ELEMENTARY CONSTRUCTIONS CLEAR.

FL	ENGLISH EQUIVALENT
(BE HAT)	THIS IS A HAT
(BE HAT[OF BOY])	THIS IS THE BOY'S HAT
(BE HAT[OF (BOY)])	THIS IS HIS HAT
(Q BE BOOK HERE)	IS THE BOOK HERE Q
(BE (ON HAT TABLE))	THE HAT IS ON THE TABLE
(BE (ON TABLE HAT))	THE TABLE IS ON THE HAT
(BE (ON HAT[OF (BOY)] TABLE))	HIS HAT IS ON THE TABLE
(BE (IN (AND HAT BOOK) DRAWER))	THE HAT AND THE BOOK ARE IN THE DRAWER
(BE (IN ((AND HAT BOOK)) (DRAWER)))	THEY ARE IN IT

INPUT TO ZBIE IS IN THE FORM OF IPLV LIST STRUCTURES EQUIVALENT TO THE ABOVE NOTATION. THE MOST IMPORTANT FEATURES OF FL SEEM TO BE ITS UNIFORMITY AND STRUCTURE : FL SENTENCES ARE USUALLY TREES, NOT STRINGS.

TO MAKE EXPLICIT THE STRUCTURE OF THE TREES IN FL, IT IS SUFFICIENT TO DEFINE A RECOGNIZER FOR THE TERMINAL NODES OF THE TREES. THE FOLLOWING ARE TERMINAL NODES:

- AN ATOMIC SYMBOL (I.E. AN IPL-V REGIONAL)

  EXAMPLE: BE, TABLE, BOY, 2.

- (<ATOMIC SYMBOL>)

  EXAMPLE: (BOY).

- (SPEAKING <ANY FL CONSTRUCT>)

  EXAMPLE: (SPEAKING BOY).

- (SPOKEN <ANY FL CONSTRUCT>)

  EXAMPLE: (SPOKEN BOY(NUMB 2)), WHERE NUMB = NUMBER.

- (<ATOMIC SYMBOL>(NUMB <ATOMIC SYMBOL>))

  EXAMPLE: (MAN(NUMB 2)).

- <ATOMIC SYMBOL>(NUMB PLUR), WHERE PLUR = PLURAL.

  EXAMPLE: BOY(NUMB PLUR).

THE TERMINAL NODES IN FL ARE CALLED FL UNITS. ALL OTHER CONSTRUCTS IN FL ARE FL COMPLEX STRUCTURES. THE PROGRAM 'UNDERSTANDS' FL TO THE EXTENT THAT IT RECOGNIZES THE FL UNITS OF AN FL STRUCTURE.

(IN THE IPL-V IMPLEMENTATION, THE SQUARE-BRACKETED DESCRIPTION LIST ACTUALLY OCCURS ABOVE THE FL UNIT THAT IS DESCRIBED, SO THAT HAT(OF BOY) LOOKS LIKE ((OF BOY) HAT) WHEN CONSIDERED PURELY AS A LIST STRUCTURE. FOR AN EXAMPLE, SEE THE IPL-V DESCRIPTION OF SENTENCE 3 IN APPENDIX B. THE ORDER 'DESCRIPTION LIST - DESCRIBED FL UNIT', IS MAINTAINED IN THE PATTERN STRUCTURES, TO BE DESCRIBED. FOR AN EXAMPLE, SEE CHAPTER III, SENTENCE 12.)

AT THIS STAGE FL IS A TOOL, TO BE MODIFIED IF NECESSARY. WE MAKE NO CLAIM THAT IT IS 'THE' REPRESENTATION, OR THAT IT IS UNIVERSAL. IT IS DOUBTFUL THAT THE PICTURE LANGUAGE IS UNIVERSAL. FOR INSTANCE, IN 'GERMAN THROUGH PICTURES', P. 239, A GERMAN BOY PLAYS BASEBALL.

8. THE PROGRAM'S INTERNAL REPRESENTATION.

AT RUN-TIME, ZBIE BUILDS AND THEN USES CERTAIN MEMORY STRUCTURES WHICH WILL NOW BE DESCRIBED.

1 - THE PATTERN.

THE MAIN WORKING STRUCTURE IS THE PATTERN, WHICH IS USED TO MATCH FL STRUCTURES, AND THEN, USING THE PATTERN'S TRANSLATION RULE, TO TRANSLATE THE FL STRUCTURES INTO NL. THERE ARE TWO TYPES OF PATTERNS, DIFFERENTIATED BY A MARKER ON THE DESCRIPTION LIST (D.L.) OF THE PATTERN. A TOP PATTERN IS USED TO MATCH FL SENTENCES; A SUBPATTERN TO MATCH FL COMPLEX SUBSTRUCTURES. A PATTERN IS AN ORDERED LIST OF PAIRS; EACH PAIR CONSISTS OF THE NAME OF A SET AND AN EXTRACTOR. ON THE D.L. OF THE PATTERN IS THE TRANSLATION RULE OF THE PATTERN, AND OTHER INFORMATION.

A MORE FORMAL DESCRIPTION OF A PATTERN CAN BE GIVEN IN B.N.F.:

```
<PATTERN>          ::= <P-LIST><DESCRIPTION LIST>
<P-LIST>           ::= <SET NAME><EXTRACTOR> |
                       <SET NAME><EXTRACTOR><P-LIST>
<DESCRIPTION LIST>::= <ATTRIBUTE><VALUE>|
                       <ATTRIBUTE><VALUE><DESCRIPTION LIST>
```

<SET NAME> ::= A1|A2|A3 ...ETC...

<EXTRACTOR> ::= Y1|Y2|Y3 ...ETC...

<ATTRIBUTE> ::= <IPL-V REGIONAL OR INTERNAL SYMBOL>

<VALUE> ::= <IPL-V LIST STRUCTURE>

IN IPL-V, A PATTERN IS A SIMPLE DESCRIBABLE LIST.

THE ELEMENTS OF A SET ARE FL UNITS OR (RECURSIVELY) PATTERNS, WHICH ARE THEN REFERRED TO AS SUBPATTERNS. A SET CAN ALSO BE EMPTY.

THE TRANSLATION RULE OF A PATTERN IS A FUNCTION OF THE EXTRACTORS OF THE PATTERN. EXAMPLE:

<P2>	<9-1 = 26178>	<9-2 = 26170>	<9-3 = 26200>
02 9-1	04 0	04 0	04 0
00 A4	00 T0	00 Y1	00 Y1
00 Y1	02 9-2	00 Y2	00 Y2
00 A3	00 D12		
00 Y2	02 9-3		
	00 D13		
	04 26180		

THE EXAMPLES ARE TAKEN FROM CHAPTER III. P2 IS THE NAME OF THE PATTERN; ITS P-LIST IS (A4 Y1 A3 Y2); ITS DESCRIPTION LIST IS THE LIST STRUCTURE 9-1. THE SETS ARE A4 AND A3; THE EXTRACTORS ARE Y1 AND Y2. THE TRANSLATION RULE OF THE PATTERN IS T0(P2)=9-2, THE LIST (Y1 Y2).

ALL OTHER ATTRIBUTE - VALUE PAIRS ON THE D.L. OF THE

PATTERNS MAY BE DISREGARDED AT THE PRESENT.

TO UNDERSTAND THE FUNCTION OF THE EXTRACTORS, LET US DESCRIBE THE PROCESS WHICH MATCHES AN FL STRUCTURE TO A PATTERN.

2 - ELEMENTARY DESCRIPTION OF THE MATCHING ROUTINE.

LET US ASSUME THAT WE WANT TO MATCH THE FL SENTENCE (BE BOY) TO P2. WE GO DOWN THE SENTENCE AND THE P-LIST OF THE PATTERN IN PARALLEL. WE CHECK WHETHER THE FIRST ELEMENT OF THE FL SENTENCE (HERE, 'BE') IS A MEMBER OF THE FIRST SET ON THE P-LIST OF THE PATTERN (HERE, A4). IN OUR CASE, 'BE' IS A MEMBER OF A4, AND WE SAY THAT 'BE' WAS (SUCCESSFULLY) MATCHED TO A4. 'BE' IS THEN PLACED ON THE DESCRIPTION LIST OF Y1; Y1 HAS 'EXTRACTED' THE ELEMENT OF THE FL SENTENCE WHICH WAS MATCHED TO A4. THE MATCHING ROUTINE LOOPS BACK AND TESTS WHETHER THE SECOND ELEMENT OF THE FL SENTENCE (HERE, 'BOY') IS A MEMBER OF THE SECOND SET ON THE P-LIST OF THE PATTERN (HERE, A3), ETC...

WHEN WE WANT TO TRANSLATE THE MATCHED FL SENTENCE INTO NL, WE USE THE TRANSLATION RULE OF THE PATTERN. THE TRANSLATION RULE OF P2 IS TO(P2)= (Y1 Y2). THE MEANING OF THE TRANSLATION RULE IS AS FOLLOWS: TAKE THE ELEMENT EXTRACTED BY Y1, 'BE', TRANSLATE IT IN THE APPROPRIATE CONTEXT (HERE, THE CONTEXT CONSISTS OF THE SET A4, OF WHICH 'BE' WAS A MEMBER, AND OF THE PATTERN P2), THEN FOLLOW THIS TRANSLATION ('E2TO') BY THE TRANSLATION OF THE ELEMENT EXTRACTED BY Y2, 'BOY', IN THE PROPER CONTEXT (HERE, THE SET A3 AND THE PATTERN P2). THE SECOND TRANSLATION IS 'MAL1YIK',

SO THAT THE TOTAL TRANSLATION IS 'E2TO MAL1YIK'. IF WE CANNOT FIND A TRANSLATION FOR 'BOY', WE INSERT A 'Z' IN THE TOTAL TRANSLATION, WHICH WOULD BECOME: 'E2TO Z'.

IT MAY HAPPEN THAT, FOR EXAMPLE, THE SECOND ELEMENT OF FL IS NOT AN FL UNIT, (SEE SENTENCE 12, CHAPTER III). THE MATCHING ROUTINE THEN TRIES, RECURSIVELY, TO FIND IN A3 A SUBPATTERN WHICH CAN MATCH THE SECOND ELEMENT OF FL.

THE MATCHING ROUTINE USES ONLY SET INCLUSION TESTS, BUT USES THEM RECURSIVELY. IT IS SEEN THAT A NECESSARY CONDITION FOR MATCH IS THAT THE LENGTH OF THE FL STRUCTURE AT ITS TOP LEVEL BE EQUAL TO THE NUMBER OF PAIRS (SET, EXTRACTOR) IN THE PATTERN. THE MATCHING ROUTINE WILL BE CONSIDERED IN DETAIL IN CHAPTER II, PART C.

3 - THE TRANSLATION RULE.

THE TRANSLATION RULE TO(PATTERN) IS A FUNCTION FROM THE EXTRACTORS OF THE PATTERN INTO NL AUGMENTED BY Z'S, WHERE A Z INDICATES THAT SOMETHING WAS NOT TRANSLATED. A FEW EXAMPLES FOLLOW:

A) LINEAR ARRANGEMENT OF THE EXTRACTORS:

<P37>	<9-1 = 26506>	<9-2 = 27334>	<9-3 = 2714 >
02 9-1	04 0	04 0	04 0
00 A12	00 T0	00 Y96	00 Y96
00 Y97	02 9-2	00 Y97	00 Y97
00 A2	00 D12	00 Y95	00 Y95

00 Y96
00 A13
00 Y95

02 9-3
00 D13
04 27144
00 D9
00 A14
00 V0
00 J4

THE TRANSLATION RULE OF P37 IS T0(P37) = (Y96 Y97 Y95), AND IS TO BE READ AS FOLLOWS: LOOK UP IN THE VOCABULARY, IN THE PROPER CONTEXT (HERE, OF THE SET A2 AND THE PATTERN P37), THE TRANSLATION OF THE PART OF THE FL STRUCTURE THAT WAS MATCHED TO A2; THEN FOLLOW THIS TRANSLATION BY THE TRANSLATION, IN THE PROPER CONTEXT, OF THE PART OF THE FL STRUCTURE THAT WAS MATCHED TO A12, ETC...

B) LINEAR ARRANGEMENT OF SOME OF THE EXTRACTORS.

<P1>	<9-1 = 25418>	<9-2 = 25752>	<9-3 = 25293>
02 9-1	04 0	04 0	04 0
00 A1	00 P0	00 P0	00 Y2
00 Y1	02 9-2		00 Y3
00 A2	00 Y0		
00 Y2	00 J3		
00 A3	00 J4		
00 Y3	00 J3		
	00 T0		
	02 9-3		

THE TRANSLATION RULE OF P1 IS T0(P1) = (Y2 Y3). IT IS TO BE READ AS IN CASE A) ABOVE. NOTICE THAT THE EXTRACTOR Y1 IS MISSING IN THE TRANSLATION RULE. SUCH A RULE CAN BE USED WHEN SOME FL PART IS NOT EXPRESSED IN NL.

C) GROUPING OF SOME EXTRACTORS.

<P0>	<9-1 = 25418>	<9-2 = 25228>	<9-3 = 25524>
02 9-1	04 0	04 0	04 0
00 A1	00 T0	02 9-3	00 Y1
00 Y1	02 9-2	00 Y3	00 Y2
00 A2			
00 Y2			
00 A3			
00 Y3			

THE TRANSLATION RULE IS TO BE READ AS FOLLOWS: TAKE THE FL STRUCTURE THAT WAS MATCHED TO A1, FOLLOW THIS STRUCTURE BY THE FL STRUCTURE THAT WAS MATCHED TO A2; THEN LOOK UP THIS FL COMPLEX STRUCTURE IN THE VOCABULARY IN THE PROPER CONTEXT (HERE, OF THE PATTERN P0). THE NL STRING OBTAINED IS FOLLOWED BY THE TRANSLATION OF THE ELEMENT EXTRACTED BY Y3, AS ABOVE, TO GIVE THE TRANSLATION OF THE STRUCTURE MATCHED TO P0.

EXAMPLE: (CHAPTER III, SENTENCE 1.) IF WE MATCH (BE (MAN) HERE) TO P0, WE SHALL LOOK UP THE FL COMPLEX (BE (MAN)) IN THE VOCABULARY, AND FOLLOW THE TRANSLATION OF (BE (MAN)) BY THE TRANSLATION OF 'HERE' (IN THE PROPER CONTEXT).

THE ABOVE THREE FUNCTIONS ARE NOW USED BY ZBIE. NOTE THAT SINCE THE FL STRUCTURE MATCHED TO A1, SAY, CAN BE A COMPLEX FL PART, A RECURSIVE TRANSLATION LOOK-UP IS IMPLICIT IN THE TRANSLATION RULES. FOR A DETAILED EXAMPLE, SEE CHAPTER III, SENTENCE 12.

THE TRANSLATION RULE FUNCTIONS CAN BE GENERALIZED IMMEDIATELY. TWO EXAMPLES OF DIFFERENT TRANSLATION RULES FOLLOW, AND MANY MORE COULD BE DREAMED UP. (WE ASSUME A PATTERN WITH EXTRACTORS Y20 AND Y21):

D) INTRODUCTION OF CONSTANTS.

```
  92  0
      Y21
      NLI
      Y20 0
```

WHERE NLI IS SOME STRING IN NL. SUCH A RULE COULD BE USED WHEN SOME EXPRESSION IN NL HAS IDIOMATIC FILLERS.

E) DISJOINT PARTS.

```
  92  0
      FIRST(Y21)
      Y20
      SECOND(Y21)   0
```

WHERE FIRST AND SECOND ARE FUNCTIONS (WHICH MUST BE DEFINED) ON THE TRANSLATION OF THE FL STRUCTURE WHICH WAS EXTRACTED BY Y21. SUCH A RULE COULD BE USED TO HANDLE SEPARABLE GERMAN VERBS.

4 - THE IN-CONTEXT VOCABULARY.

THE VOCABULARY OF ZBIE HAS THE TWO FOLLOWING FORMS:

A) FL UNIT=FL(I), SET=A(J), PATTERN=P(K), NL STRING=NL(L).
TO BE INTERPRETED: THE TRANSLATION OF THE FL UNIT FL(I) WHEN A MEMBER OF THE SET A(J) AND IN THE CONTEXT OF THE PATTERN P(K) IS THE STRING OF NL WORDS NL(L).

B) FL COMPLEX=FL(I), PATTERN=P(J), NL STRING=NL(K)
TO BE INTERPRETED: THE TRANSLATION OF THE FL COMPLEX FL(I) IN THE CONTEXT OF THE PATTERN P(J) IS THE STRING NL(K) IN NL.

NOTE THAT WE CANNOT CONCLUDE THAT SOME FL UNIT FL(I) HAS A TRANSLATION IN THE CONTEXT A(J), P(K) FROM THE KNOWLEDGE THAT FL(I) IS A MEMBER OF A(J). WE MAY HAVE A TRANSLATION IN THE CONTEXT OF SOME OTHER PATTERN P(K'), OR, FOR THAT MATTER, NONE AT ALL.

C. THE PROGRAM'S ORGANIZATION.

THE CONTROL STRUCTURE OF ZBIE CONSISTS OF TWO MODES. THERE IS FIRST AN INITIALIZATION OF THE INTERNAL STRUCTURE, WHEN A FIRST PATTERN IS CONSTRUCTED. THE CONTROL THEN PASSES TO A SECOND MODE WHEN SITUATIONS ARE BROUGHT IN ONE BY ONE AND PROCESSED.

MODE 1. INITIALIZATION.

TWO SITUATIONS ARE PRESENTED TO ZBIE, REPRESENTED IN FL AND EXPRESSED IN NL. THE SITUATIONS MUST BE SUFFICIENTLY SIMILAR, SO THAT BY COMPARING THEM ZBIE CAN DEDUCE ITS FIRST PATTERN.

MORE PRECISELY, ZBIE EXPECTS: THAT THE TWO FL SENTENCES WILL BE OF DEPTH 0, I.E. HAVE NO FL COMPLEX SUBSTRUCTURE; THAT THEY WILL HAVE EXACTLY ONE ELEMENT DIFFERENT AND IN THE SAME POSITION; AND THAT THE TWO CORRESPONDING NL SENTENCES WILL HAVE EXACTLY ONE ELEMENT DIFFERENT AND IN THE SAME POSITION, WHICH MUST BE AT THE BEGINNING OR AT THE END OF (EITHER) NL SENTENCE. THE DISTINCT ELEMENTS IN FL AND NL ARE THEN ASSUMED TO CORRESPOND TO EACH OTHER; THE COMMON PARTS IN FL AND NL ARE ALSO ASSUMED TO CORRESPOND TO EACH OTHER AND A FIRST PATTERN P0 IS SET UP, WITH ITS SETS AND TRANSLATION RULE; THE IN-CONTEXT VOCABULARY IS INITIATED.

MODE 2. SINGLE SENTENCE ANALYSIS.

AFTER INITIALIZATION, ZBIE OPERATES IN THE FOLLOWING MODE. FIRST, THE PREVIOUSLY PROCESSED FL AND NL SENTENCES ARE ERASED. THEN, THE DESCRIPTION IN FL OF A SITUATION IS READ IN, TOGETHER WITH ITS EXPRESSION IN NL WHICH IS STORED FOR LATER USE. THE FL SENTENCE IS THEN PROCESSED FOLLOWING THE BASIC FLOW CHART BELOW (WRITTEN IN AN ALGOL-LIKE LANGUAGE, '<' AND '>' DELIMIT BLOCKS):

SINGLE SENTENCE PROCESSOR.

```
FOR ALL TOP PATTERNS (LAST CREATED, FIRST CONSIDERED) DO
  <MATCH FL TO TOP PATTERN;
  COMMENT: THE MATCHING ROUTINE IS DESCRIBED IN PART 1 BELOW;
  IF TOTAL MATCH THEN
    <TRANSLATE;
    IF TRANSLATION HAS NO UNKNOWNS THEN <COMPARE TO INPUT NL;
      IF TRANSLATION = INPUT NL THEN <EXIT AND READ IN THE NEXT
      SITUATION> ELSE GO TO ERROR RECOVERY>
    ELSE IF TRANSLATION IS CONSISTENT WITH INPUT NL THEN STORE
    PATTERN LIST IN PATTERN LIST HOLDER TO PROCESS ELSE DO
    NOTHING>
    COMMENT: THE CONSISTENCY TEST IS DESCRIBED IN PART 2;
  ELSE STORE PATTERN LIST IN PATTERN LIST HOLDER>
PROCESS ELEMENTS ON THE PATTERN LIST HOLDER
COMMENT: PROCESSING THE PATTERN LISTS IS DESCRIBED IN PART 3;
  <IF PROCESSING SUCCESSFUL THEN <EXIT AND READ IN THE NEXT
```

SITUATION>>

COMMENT : EVERYTHING FAILED SO FAR;

CREATE A NEW TOP PATTERN FOR THE SITUATION.

COMMENT: THE PATTERN CREATING ROUTINE IS DESCRIBED IN PART 4;

BEFORE DESCRIBING HOW THE PATTERN LISTS ARE PROCESSED, LET US EXPLAIN SOME OF THE TERMS USED.

WE SAW IN CHAPTER II, PART 8 HOW AN FL SENTENCE WAS MATCHED TO A PATTERN. ONE OF THE MAIN ROUTINES OF ZBIE IS THE MATCHING ROUTINE, WHICH OBTAINS ALL POSSIBLE TOTAL AND PARTIAL MATCHES OF PATTERNS (AND SUBPATTERNS) TO AN FL SENTENCE. LET US CONSIDER THE MATCHING MECHANISM AGAIN.

PART 1 - THE MATCHING ROUTINE.

AN FL SENTENCE IS A TREE STRUCTURE; THE TERMINAL NODES ARE THE FL UNITS. THE SETS OF THE P-LIST OF A TOP PATTERN CAN BE VIEWED AS THE ZERO-TH LEVEL OF A TREE, THE PATTERN TREE. THE SETS THEMSELVES CAN CONTAIN (SUB)PATTERNS; WE CAN THINK OF EACH SUBPATTERN AS INITIATING A NEW BRANCH OF THE PATTERN TREE. AN INSTANCE OF A SET WILL BE CONSIDERED AS A TERMINAL NODE OF THE PATTERN TREE IF WE SELECT AN FL UNIT OF THE SET. IF WE SELECT A SUBPATTERN OF THE SET, THE INSTANCE OF THE SET WILL BE CONSIDERED AS A NON-TERMINAL NODE, AND THE SUBPATTERN WILL INITIATE A NEW SUBTREE. SINCE A SUBPATTERN SP1 MAY HAVE ON ITS P-LIST THE NAME OF A SET WHICH CONTAINS A (SUB)PATTERN P2 SUCH THAT P2 IS AN ANCESTOR OF SP1, A TOP PATTERN CAN BE CONSIDERED AS THE HEAD OF

AN INFINITY OF TREES OF ARBITRARILY LARGE DEPTH. HOWEVER, SINCE ANY FL SENTENCE IS A FINITE TREE, NO PATTERN TREES OF A DEPTH SUPERIOR TO THE DEPTH OF THE FL SENTENCE NEED BE CONSIDERED.

THE PATTERN TREES FORMED BY THE PATTERNS RESEMBLE A DISCRIMINATION NET. A NET OF DEPTH N CANNOT DISTINGUISH BETWEEN TWO TREES OF DEPTH LARGER THAN N WHICH ARE IDENTICAL UP TO AND INCLUDING DEPTH N. HOWEVER, IN SOME CASES, N DIFFERENT PATTERNS COULD DISTINGUISH TWO SUCH TREES THANKS TO THE RECURSIVE FEATURE OF THE PATTERNS. THE PATTERN TREES ARE, THEREFORE, MORE POWERFUL THAN A SIMPLE DISCRIMINATION NET.

WE CAN VIEW THE MATCHING PROCESS AS A TEST OF WHETHER A PATTERN TREE IS CLOSE TO BEING ISOMORPHIC TO THE FL SENTENCE TREE. AN FL SENTENCE TREE IS ISOMORPHIC TO A PATTERN TREE IF THE TREES CAN BE SUPERIMPOSED, AND THE TERMINAL NODES (FL UNITS) OF THE FL SENTENCE ARE RESPECTIVELY IDENTICAL TO THE TERMINAL NODES (FL UNITS OF TERMINAL SETS) OF THE PATTERN TREE. USUALLY WE SHALL HAVE NO ISOMORPHISM; WE THEN LOOK FOR AS GOOD A FIT OF A PATTERN TREE TO THE FL SENTENCE AS WE CAN FIND.

DISREGARDING THE BOOKKEEPING CHORES INVOLVED IN BACKTRACKING AND RECURSION, IT IS ENOUGH TO CONSIDER HOW AN FL SUBTREE IS MATCHED TO AN ELEMENT OF A SET, NODE OF THE PATTERN TREE. BY CONSIDERING THE TOP PATTERNS AS ELEMENTS OF A SET, NO GENERALITY IS LOST. WE TREAT THE SPECIAL CASE OF AN EMPTY SET BY POSTULATING AN EMPTY ELEMENT FOR SUCH A SET.

PART 1.1 - OUTLINE OF THE MATCHING ROUTINE.

```
MATCH FL (SUB)TREE TO ELEMENT OF SET;
COMMENT: QUICK EXITS;
ELSET := ELEMENT OF SET;
FLTREE:= FL SUBTREE;
(IF ELSET IS AN FL WHOLE OR IF ELSET IS EMPTY) THEN <'NO MATCH';
GO TO EXIT>;

COMMENT: ELSET IS A (SUB)PATTERN;
IF LENGTH (P-LIST OF ELSET) =/= 2 * LENGTH OF FLTREE
  THEN <'NO MATCH'; GO TO EXIT>;
COMMENT: INITIALIZE LOOP;
MISTAKE COUNTER := 0;
COMMENT:  THE  MISTAKE COUNTER IS ASSOCIATED WITH THIS PARTICULAR
LEVEL, AS ARE THE OTHER IDENTIFIERS;

I := 0;
LOOP:
IF MISTAKE COUNTER > 1 THEN <'NO MATCH'; GO TO EXIT>;
I := I+1;
IF I-TH SON OF FLTREE DOES NOT EXIST THEN GO TO EXIT;
IFLTREE := I-TH SON OF FLTREE;
IELSET  := I-TH SET OF P-LIST OF ELSET;
IF IFLTREE IS AN FL WHOLE THEN
  <IF IFLTREE IS A MEMBER OF IELSET THEN
    GO TO LOOP
```

```
  ELSE <MISTAKE COUNTER := MISTAKE COUNTER + 1; GO TO LOOP>>
ELSE <MATCH IFLTREE TO THE ELEMENTS OF IELSET;
COMMENT: THE MATCH IS TRIED SUCCESSIVELY ON ALL THE ELEMENTS OF
IELSET.  THANKS TO THE HIDDEN BOOKKEEPING, WE ONLY HAVE TO
CONSIDER THE MATCH FOR ONE ELEMENT;

IF RETURN 'NO MATCH' THEN
  <MISTAKE COUNTER := MISTAKE COUNTER + 1; GO TO LOOP>
  ELSE GO TO LOOP>
```

IT IS SEEN THAT, BASICALLY, UP TO ONE 'MISTAKE' (AS DEFINED BY THE PROGRAM) IS ALLOWED AT A GIVEN LEVEL IN A SUBTREE.

PART 1.2 - USE OF THE MATCHING ROUTINE.

THE MATCHING ROUTINE FINDS PATTERN LISTS, WHICH IT STORES IN A PATTERN LIST HOLDER IN DECREASING ORDER OF MATCH-DEPTH. A PATTERN LIST IS A LIST OF PATTERNS, HEADED BY A TOP PATTERN AND CONTAINING OTHER (OR POSSIBLY NO) SUBPATTERNS WHICH WERE MATCHED TO SUBSTRUCTURES OF THE FL SENTENCE. WITH THE PATTERN LIST IS ASSOCIATED A CORRESPONDING LIST OF FL COMPLEX STRUCTURES IN ONE-TO-ONE CORRESPONDENCE WITH THE SUBPATTERNS THEY MATCHED. THE MATCHING ROUTINE ALSO RECORDS THE DEEPEST LEVEL IN THE FL SENTENCE TO WHICH THE MATCH WAS CARRIED, THE MATCH-DEPTH, AND WHETHER THE MATCH WAS TOTAL OR PARTIAL. A MATCH IS PARTIAL IF AT ANY POINT, DURING THE MATCH OF A PATTERN TREE TO THE FL SENTENCE, A MISTAKE COUNTER WAS INCREMENTED. THE MATCH-DEPTH IS A MEASURE OF HOW GOOD THE MATCH OF THE FL SENTENCE TO THE PATTERN LIST HAS

BEEN, AND ZBIE LOOKS AT THE BEST MATCHES FIRST.

WHEN THE TRANSLATION OF A PATTERN LIST IS ATTEMPTED, IT MAY HAPPEN THAT SOME FL STRUCTURE CANNOT BE TRANSLATED. FOR EXAMPLE, THE STRUCTURE MAY BE AN FL UNIT WITH NO TRANSLATION IN THE CONTEXT CONSIDERED, OR IT IS AN FL COMPLEX PART WHICH, FOR THE PARTICULAR PATTERN LIST CONSIDERED, HAS NO CORRESPONDING SUBPATTERN TO WHICH IT WAS MATCHED. WE THEN INSERT A Z (Z0, Z1, Z2, ...) IN THE TRANSLATION. THE Z IS CONSIDERED BY ZBIE TO BE THE RESULT OF AN UNFULFILLED EXPECTATION, AND ZBIE CAPITALIZES ON THESE EXPECTATIONS FOR LEARNING. ZBIE CHECKS WHETHER THE TRANSLATION OBTAINED IS CONSISTENT WITH THE INPUT NL SENTENCE.

PART 2 - THE CONSISTENCY TEST.

WE CAN IMAGINE THAT THE TRANSLATION AND THE INPUT NL SENTENCE ARE PUT SIDE BY SIDE. WE SHALL HAVE CONSISTENCY IF WE CAN REPLACE THE Z'S OF THE TRANSLATION BY NON-EMPTY STRINGS IN NL IN A UNIQUE NON-AMBIGUOUS WAY SO THAT THE TRANSLATION BECOMES IDENTICAL TO THE INPUT. WHETHER THIS CAN BE DONE OR NOT IS OFTEN TRIVIAL IF NO TWO Z'S ARE ADJACENT. WHEN TWO OR MORE Z'S ARE ADJACENT, ZBIE DOES NOT GIVE UP BUT REPLACES THE FIRST Z BY AN APPROPRIATE GOOD GUESS (IF AVAILABLE, SEE BELOW). THE VARIOUS GUESSES THAT WILL BE TRIED ARE PRINTED. IF PROGRESS IS MADE TOWARDS CONSISTENCY, THE GUESS IS ADOPTED (WE DO NOT HAVE COMPLETE BACKTRACKING HERE), AND WE CONTINUE. OTHERWISE, GUESSES ARE TRIED FOR THE SECOND Z. IF THIS LAST RESORT FAILS, THE TRANSLATION AND THE INPUT NL SENTENCE ARE NOT CONSISTENT.

HERE WE SEE AN EXAMPLE OF THE CAUTION ZBIE USES IN LEARNING. IF WE WANT Z Z1 TO CORRESPOND TO THE NL STRING NL1 NL2 NL3, WE CAN MAKE THE CORRESPONDENCE IN TWO WAYS (THE Z'S MUST BE MATCHED TO NON-EMPTY STRINGS IN NL):

Z←NL1, Z1←NL2 NL3;

Z←NL1 NL2, Z1←NL3;

HOWEVER, IF IT IS A GOOD GUESS TO ASSUME THAT Z←NL1, THEN WE CAN LET Z1←NL2 NL3.

AN EXAMPLE FROM CHAPTER III (SENTENCE 4) WILL ILLUSTRATE THE POINT. WE ARE TESTING THE CONSISTENCY OF (Z Z1) AND (TI2 $ZDES_1$). (Z Z1) IS THE TRANSLATION OF (BE (SPOKEN BOY) HERE). 'Z' COMES FROM THE UNKNOWN TRANSLATION OF (SPOKEN BOY) AND '71' FROM THE UNKNOWN TRANSLATION OF 'HERE'. WITH NO ADDITIONAL INFORMATION, ZBIE REFUSES TO MAKE A CORRESPONDENCE BETWEEN THE Z'S AND THE NL WORDS, AND WOULD FIND THE SENTENCE 'TOO HARD'. (STRICTLY, WE SHOULD LET Z←TI2, Z1←ZDES1 AS Z'S CORRESPOND TO NON-EMPTY STRINGS, BUT ZBIE IS EVEN MORE CAUTIOUS.) HOWEVER, IT IS A 'GOOD GUESS', IN THIS CONTEXT, TO ASSUME THAT THE TRANSLATION OF 'HERE' IS 'ZDES1', SO THAT, AS A RESULT, Z←TI2.

IF SOME PATTERN LIST COMPLETELY MATCHES AN FL SENTENCE AND WE OBTAIN A TOTAL TRANSLATION (WITHOUT ANY Z'S) WHICH IS NOT IDENTICAL TO THE INPUT NL, SEVERAL POSSIBILITIES ARISE:

1) THERE WAS A MISTAKE IN THE INPUT (THAT HAS HAPPENED).

2) THE TRANSLATION AND THE INPUT ARE TWO DIFFERENT WAYS OF

EXPRESSING THE SAME SITUATION. HERE A TEACHER (AND PREFERABLY A TIME SHARING SITUATION) IS NEEDED TO CONVEY THE INFORMATION TO ZBIE.

3) ZBIE MADE SOME ERROR SOMEWHERE AND SHOULD TRY TO RECOVER FROM THIS ERROR. GOOD ERROR RECOVERY IS A VERY HARD PROBLEM WHICH WILL BE MENTIONED LATER. AT THIS STAGE, ZBIE POSSESSES NO ERROR RECOVERY MECHANISM. IT WAS FELT THAT TRYING TO AVOID ERRORS IS A MORE FRUITFUL APPROACH THAN TRYING TO RECOVER FROM THEM. THE ERROR AVOIDING MECHANISM IS POWERFUL ENOUGH SO THAT ZBIE ACTUALLY MAKES NO ERRORS WHEN TESTED ON THE SIMPLE SENTENCES GIVEN AS EXAMPLES.

PART 3 - PROCESSING THE PATTERN LISTS.

WE NOW RETURN TO THE BASIC FLOW-CHART. THE PROCESSING IS SLIGHTLY DIFFERENT FOR FL SENTENCES WITH FL COMPLEX PARTS THAN FOR LINEAR FL SENTENCES (OF DEPTH 0). LET US DESCRIBE PROCESSING THE FORMER SENTENCES.

PART 3.1 - PATTERN LISTS OF FL SENTENCES WITH COMPLEX PARTS.

COMMENT : PROCESS PATTERN LISTS;
FOR ALL PATTERN LISTS OF THE DEEPEST LEVEL (THEN DEEPEST LEVEL - 1, ETC...) DO
  <TRANSLATE PATTERN LISTS;
  KEEP ONLY PATTERN LISTS WHICH HAVE A TRANSLATION CONSISTENT WITH THE INPUT NL;
  PROCESS PATTERN LISTS WITH CONSISTENT TRANSLATION, STARTING

WITH THE PATTERN LISTS THAT HAVE A TRANSLATION WITH THE GREATEST NUMBER OF COMMON ELEMENTS WITH THE INPUT NL SENTENCE>

IN OTHER WORDS, ZBIE DOES A TRANSLATION 'IN PARALLEL' OF ALL PATTERN LISTS THAT MATCHED THE INPUT FL SENTENCE TO A GIVEN DEPTH, STARTING WITH THE MAXIMUM DEPTH (BEST MATCHES) FIRST. ZBIE THEN DISCARDS THE PATTERN LISTS WHICH DID NOT GIVE A CONSISTENT TRANSLATION AND STARTS PROCESSING THE PATTERN LISTS WITH THE BEST FIT TO THE INPUT, AS MEASURED BY A SET INTERSECTION WITH THE INPUT NL SENTENCE.

3.1.1 - PROCESSING THE Z'S.

PROCESSING SUCH A TRANSLATION IS EQUIVALENT TO PROCESSING THE Z'S. EACH Z TAKES THE PLACE OF SOME UNTRANSLATED FL STRUCTURE FLZ AND, THROUGH THE CONSISTENCY TEST (PART 2 ABOVE), THE Z IS TO BE REPLACED BY A NON-EMPTY NL STRING NLZ. FROM THE Z WE CAN ALSO OBTAIN INFORMATION SUCH AS: TO WHICH SET, AZ, DID WE TRY TO MATCH FLZ; WHICH WAS THE PATTERN, PZ, TO WHICH THE FATHER OF FLZ (IN THE FL SENTENCE TREE) WAS BEING MATCHED.

IF FLZ IS A UNIT, WE INSERT FLZ IN AZ AND SET UP THE IN-CONTEXT VOCABULARY FLZ AZ PZ NLZ.

IF FLZ IS A COMPLEX FL, WE TRY TO CREATE A LIST OF SUBPATTERNS TO MATCH FLZ TO NLZ (SEE BELOW). IF SUCCESSFUL, WE WOULD LIKE TO INSERT THE TOP SUBPATTERN OF THE LIST AT THE TOP OF AZ.

BEFORE INSERTING, WE CHECK TO MAKE SURE THAT THE SUBPATTERN LIST WILL NOT CAUSE AMBIGUITIES. THIS CAN HAPPEN IF THERE IS ALREADY IN AZ A SUBPATTERN LIST SP1 OF WHICH THE NEW SUBPATTERN LIST SP2 IS A HOMOMORPHIC IMAGE, I.E. SETS OF SP2 ARE SUBSETS OF THE CORRESPONDING SETS OF SP1 AND THE TRANSLATIONS OF THE SUBPATTERNS IN SP1 AND SP2 ARE APPROPRIATELY IDENTICAL. IF SUCH A CONDITION IS SATISFIED (WITH SOME MINOR ADDITIONS), NO INSERTION TAKES PLACE, 'NOT INSERTED' IS PRINTED, AND PROCESSING OF THE PATTERN LIST IS ENDED.

IF AN FL COMPLEX PART IS TO BE MATCHED TO A SET, THE BACK-TRACKING MATCHING ROUTINE WILL SEARCH FOR ALL THE SUBPATTERNS (OF THE SET) THAT WILL MATCH THE FL COMPLEX PART. IF TWO SUBPATTERNS CAN TRANSLATE THE FL COMPLEX PART (IN DIFFERENT WAYS), THEN A POTENTIAL AMBIGUITY IS INTRODUCED IN THE SYSTEM. EXAMPLE: IN CHAPTER III, SENTENCE 27, THE FL COMPLEX (MOD THIS) IS TRANSLATED AS 'E2TOT' BY SUBPATTERN P33 AND AS 'E2TA' BY SUBPATTERN P32. THE TWO SUBPATTERNS WILL CAUSE AMBIGUITY IF THEY BELONG TO THE SAME SET.

HERE IS ANOTHER EXAMPLE OF HOW ZRIE TRIES TO AVOID ERRORS THE CONTEXT IN WHICH THE SUBPATTERNS WERE ORIGINALLY BUILT WAS TOO SMALL, AND POSSIBLY LATER, BY WIDENING THE CONTEXT, SUBPATTERNS CAN BE BUILT WITHOUT RISKING AMBIGUITIES AT A LATER TIME.

NOTE THAT WE INSERT THE HEAD OF A SUBPATTERN LIST AT THE TOP

OF AZ SO THAT IF THE SAME FL SENTENCE IS PRESENTED ONCE MORE IMMEDIATELY AFTERWARDS, IT WILL BE MATCHED WITHOUT ANY BACKTRACKING USING THE SUBPATTERN(S) JUST CREATED.

PART 3.2 - PATTERN LISTS OF LINEAR FL SENTENCES.

WHEN PROCESSING LINEAR FL SENTENCES, WE ARE ONLY CONCERNED WITH TOP PATTERNS. PROCESSING ALL TOP PATTERNS WHICH GAVE A PARTIAL MATCH WOULD BE WASTEFUL, SO WE PROCESS THEM AS A STACK; LAST CREATED, FIRST CONSIDERED. HOWEVER IF A PATTERN IS FIRST CONSIDERED, AND ITS TRANSLATION IS JUST Z (NOTHING IN NL), THEN 'WAIT' IS PRINTED AND THE PATTERN IS INSERTED AT THE BOTTOM OF THE STACK FOR RECONSIDERATION LATER, IF NECESSARY. WE DO THIS BECAUSE WE MAY WELL FIND ANOTHER PATTERN WITH A TRANSLATION THAT CONTAINS SOME NL ELEMENTS AND IS CONSISTENT WITH THE INPUT.

PART 3.2.1 - MATCH-BACK.

THE MAJOR DIFFERENCE IN PROCESSING LINEAR OR NON-LINEAR FL SENTENCES IS THAT IF WE UPDATE THE VOCABULARY BY TRANSLATING AN FL COMPLEX PART IN THE CONTEXT OF A PATTERN ONLY, I.E. THE TRANSLATION RULE OF THE PATTERN WAS OF TYPE C (SEE ABOVE, CHAPTER II-8), THEN WE TRY TO SPLIT THE TRANSLATION USING EXACTLY THE SAME ROUTINES THAT WERE USED FOR THE INITIALIZATION OF THE SYSTEM. IF SUCCESSFUL, WE CREATE A NEW PATTERN WHICH WILL HAVE THE SAME SETS AND EXTRACTORS AS THE PATTERN CONSIDERED, BUT WHICH WILL HAVE A DIFFERENT (AND SIMPLER) TRANSLATION RULE. WE CALL

THIS PROCESS MATCH-BACK. FOR AN EXAMPLE, SEE CHAPTER III, SENTENCE 3; PATTERN P1 IS OBTAINED FROM PATTERN P0 BY MATCH-BACK.

SINCE WE TRY TO MATCH AN FL SENTENCE TO PATTERNS IN THE REVERSE ORDER IN WHICH THEY WERE CREATED (LAST PATTERN CREATED, FIRST TRIED), AN FL SENTENCE CAN BE MATCHED AND TRANSLATED BY A PATTERN (OBTAINED BY MATCH-BACK) WHICH IS NEWER THAN THE PATTERN THAT WOULD OTHERWISE HAVE MATCHED THE FL SENTENCE. AS A RESULT, IT CAN HAPPEN THAT, EFFECTIVELY, SOME OF THE OLDER PATTERNS ARE NEVER REACHED ANY MORE. THIS RESULT CAN BE BENEFICIAL AS SOME OF THE EARLIER PATTERNS MAY HAVE INCORPORATED MISTAKES WHICH, THEREFORE, WILL NOT BE MADE ANY MORE. WE HAVE HERE A RATHER ELEMENTARY EXAMPLE OF WHAT MAY BE CALLED 'EVOLUTIONARY LEARNING'. AN EXAMPLE IN GERMAN APPEARS IN APPENDIX B.

PART 3.2.2 - 'TRY LEARN MORE'.

IF A LINEAR FL SENTENCE HAS BEEN TRANSLATED BY A PATTERN PI, AND IF THE SENTENCE WAS COMPLETELY MATCHED BY SOME OTHER PATTERN PJ WHICH HAD BEEN CREATED AFTER PI, THEN ZBIE TAKES THE FIRST SUCH PATTERN PJ THAT WAS CONSIDERED AND TRIES TO LEARN FROM PJ, AS IF THE SENTENCE HAD NOT BEEN TRANSLATED. ZBIE PRINTS 'TRY LEARN MORE' WHEN SUCH A CASE OCCURS. IT IS THE SIMPLE 'TRY LEARN MORE' PROCESS WHICH MAKES 'EVOLUTIONARY LEARNING' POSSIBLE. WE ONLY USE THIS PROCESS FOR LINEAR FL SENTENCES. FOR COMPLEX FL SENTENCES, CONTEXT IS ESSENTIAL, AND THE CONTEXT EXISTS MORE AT THE LEVEL OF THE IN-CONTEXT TRANSLATION THAN AT THE LEVEL OF THE

SET INCLUSION TESTS.

PART 4 - THE PATTERN CREATING ROUTINE.

FINALLY, WE MUST CONSIDER ANOTHER VERY IMPORTANT ROUTINE IN ZBIE: THE PATTERN CREATING ROUTINE. DEPENDING ON AN INPUT PARAMETER, THIS ROUTINE CREATES A TOP PATTERN OR A SUBPATTERN. WE SAW ABOVE HOW THE NEED FOR NEW SUBPATTERNS ARISES. ZBIE ATTEMPTS TO CREATE A TOP PATTERN WHEN ALL PREVIOUSLY DESCRIBED PROCESSING HAS FAILED.

THE PATTERN CREATING ROUTINE IS ONE OF THE LONGEST AND MOST COMPLICATED IN THE SYSTEM. ONLY ITS MAIN PARTS WILL BE DESCRIBED. IT TRIES TO CREATE A LIST OF PATTERNS WHICH WILL MATCH THE INPUT FL COMPLEX STRUCTURE FLI TOTALLY AND GIVE AS A TRANSLATION (AFTER SUCH A MATCH) THE INPUT NL STRING NLI, MAKING USE OF ALREADY KNOWN INFORMATION.

THE ROUTINE MAKES FROM ONE TO FOUR PASSES ON FLI. EACH PASS IS FIRST PERFORMED ON ALL APPROPRIATE ELEMENTS OF FLI BEFORE THE NEXT PASS IS BEGUN. THE MAIN FEATURES OF EACH PASS ARE:

PART 4.1 - OUTLINE OF PATTERN CREATING ROUTINE.

1) FIND WHETHER AN FL UNIT (NOT A VERB) IN FLI HAS A TRANSLATION (IN SOME CONTEXT A, P) WHICH IS FOUND EXACTLY IN NLI. IF THAT IS THE CASE, AND FINALLY THE PATTERNS ARE CREATED, THEN THE SAME SET A WILL BE USED IN SAY SOME (SUB)PATTERN PJ, AND ON THE SET A WE PUT THE INFORMATION: IF WE WANT THE TRANSLATION OF AN ELEMENT IN

A IN THE CONTEXT P, IT IS A GOOD GUESS TO ASSUME THAT THE TRANSLATION IS THE SAME AS IN THE CONTEXT A, PJ AND VICE-VERSA. THIS IS A CASE OF THE 'GOOD GUESS' MENTIONED WHEN DISCUSSING THE CONSISTENCY OF A TRANSLATION WITH AN NL SENTENCE.

2) FIND WHETHER AN FL WHOLE (NOT A VERB) IN FLI HAS A TRANSLATION WHICH 'LOOKS LIKE' SOME STRING IN NLI. A TEST FOR SUCH A SIMILARITY IS MADE USING THE FIRST FEW CHARACTERS OF THE PRINT NAMES OF THE TRANSLATION AND OF THE ELEMENTS IN NLI. FOR INSTANCE, IN RUSSIAN, THE GENITIVE OF BOY LOOKS LIKE THE NOMINATIVE OF BOY, (REMEMBER THAT, BECAUSE OF THE FL STATEMENT, WE EXPECT SOMETHING WHICH HAS TO DO WITH 'BOY'). SUCH A TEST WOULD HAVE TO BE IMPROVED TO WORK FOR LANGUAGES (SUCH AS HEBREW) WHERE PREFIXES ARE ADDED TO WORDS.

3) TREAT VERBS AS IN 1). (VERBS VARY TOO MUCH TO BE RELIABLE AT FIRST).

4) IF SOME FL WHOLE IN FLI HAD PREVIOUSLY BEEN FOUND NOT TO BE TRANSLATED (WE HAD A PARTIAL TRANSLATION RULE, OF TYPE B, FOR INSTANCE), THEN ASSUME THAT AGAIN IT WILL NOT HAVE TO BE TRANSLATED.

AFTER THE APPLICATION OF A PASS TO AN ELEMENT OF FLI, A CHECK IS MADE FOR CERTAIN TERMINATING CONDITIONS: ALL FL WHOLES IN FLI USED UP, ALL NL ELEMENTS IN NLI USED UP, OR ONLY ONE FL WHOLE LEFT UNACCOUNTED FOR. THE SYSTEM ALSO CHECKS FOR POSSIBLE AMBIGUITIES. FOR EXAMPLE, SUPPOSE TWO DIFFERENT FL WHOLES IN FLI

HAVE HAD IDENTICAL TRANSLATIONS, TO BE FOUND IN NLI; THERE IS NO WAY TO MAKE A CORRESPONDENCE, AND THE PATTERN CREATING ROUTINE EXITS WITH 'TOO HARD'.

THE NEXT STEP IS TO BUILD TRANSLATION RULES FOR THE SUBPATTERNS CREATED. AT THIS STAGE, ONLY TRANSLATION RULES OF TYPES A) AND B) ARE ALLOWED. THE RULES ARE BUILT BY CONSIDERING THE NL STRING INPUT TO THE ROUTINE AND MAKING SURE THAT THE TRANSLATION OF THE PATTERNS BUILT WHEN MATCHED TO THE INPUT FL

CHAPTER III.

LEARNING RUSSIAN.

WE ARE PRESENTING HERE THE OUTPUT OF A COMPUTER RUN DURING WHICH ZBIE ATTEMPTS TO LEARN RUSSIAN. THE OUTPUT IS REPRODUCED AS OBTAINED FROM THE PRINTER. INPUTS IN IPL-V FORM HAVE BEEN DELETED EXCEPT IN ONE ILLUSTRATIVE CASE.

THE IMPORTANT ATTRIBJTES AND VALUES OF A PATTERN PJ HAVE THE FOLLOWING MEANING:

T0(PJ) = TRANSLATION RULE (A LIST STRUCTURE).
P0(PJ) = A LIST OF PATTERNS WITH A P-LIST IDENTICAL TO PJ.
V0(PJ) = J4 IF PJ IS A SUBPATTERN.
THE OTHER ATTRIBUTES ARE USED FOR BOOKKEEPING AND MAY BE DISREGARDED

COMMENTS ARE GIVEN AT THE RIGHT OF INPUTS, OR AT THE END OF THE PROCESSING OF A SENTENCE.

LOOKING AT SENTENCE COMMENTS
(BE (MAN)HERE)
ON ZDES1
(BE (MAN)THERE)

```
ON TAM                          THE FIRST TWO SENTENCES
PROCESS START                   TO START THE INITIALIZATION.
PUT INTO VOCABULARY             BE IS IN SET A1;(WOMAN) IN SET A2.
(BE (MAN ))
ON
P                               CONTEXT VOCABULARY FOR FL COMPLEX
PUT INTO VOCABULARY
HERE                            IN-CONTEXT VOCABULARY FOR FL
ZDES1                           UNIT. THE FL UNIT (IN THIS CASE,
A3                              HERE) IS INSERTED IN THE SET (IN
P                               THIS CASE, A3). P=P0.
PUT INTO VOCABULARY
THERE
TAM
A3
P
NEW PATTERN
```

<P0>	<9-1 = 25418>	<9-2 = 25228>	<9-3 = 25524>
02 9-1	04 0	04 0	04 0
00 A1	00 T0	02 9-3	00 Y1
00 Y1	02 9-2	00 Y3	00 Y2
00 A2			
00 Y2			
00 A3			
00 Y3			

COMMENT: THE INITIALIZATION PHASE IS OVER. FOR THE NEXT SENTENCE, WE SHOW THE IPL-V INPUT.

COMMENT: 'PUT INTO VOCABULARY' HAS THREE OR FOUR ARGUMENTS EXAMPLES OF BOTH CASES ARE ILLUSTRATED ABOVE:

1) THREE ARGUMENTS. THE FIRST ARGUMENT IS AN FL COMPLEX, THE SECOND ARGUMENT IS AN NL STRING, THE THIRD ARGUMENT IS A PATTERN GIVING THE CONTEXT. THE TRANSLATION OF THE FL COMPLEX IN THE CONTEXT OF THE PATTERN IS THE NL STRING. EXAMPLE: (BE(MAN)), ON, P(=P0).

2) FOUR ARGUMENTS. THE FIRST ARGUMENT IS AN FL UNIT; THE SECOND, AN NL STRING; THE THIRD, A SET AND THE FOURTH, A PATTERN. THE FL UNIT IS FIRST INSERTED IN THE SET (THE UNIT MAY ALREADY BE IN THE SET). THE TRANSLATION OF THE FL UNIT AS A MEMBER OF THE SET, IN THE CONTEXT OF THE PATTERN, IS THE NL STRING. WE REFER TO THE PAIR (SET, PATTERN) AS THE CONTEXT IN WHICH THE TRANSLATION OF THE FL UNIT IS THE NL STRING. EXAMPLES: HERE, ZDES1, A3, P(=P0): THERE, TAM, A3, P(=P0).

ALL ADDITIONS TO THE VOCABULARY ARE PRINTED AS DESCRIBED ABOVE.

```
IP INPUT DATA                          5       01
                                         X4      95     0
                                         95      0
IP BE                                            E0
IP (SPEAKING MAN)                                91
```

```
IP HERE                                    E3    0
IP                                  91     0
IP SPEAKING                                E1
IP MAN                                     E2    0
IP                                  X3     95    0
IP                                  95     0
IP A1                                      R1
IP ZDES1                                   R2    0
IP                                  5      J0
H2 = 2357  CELLS          00:00:06
LOOKING AT SENTENCE 3
(BE (SPEAKING MAN )HERE )
A1 ZDES1                    THE INPUTS ARE FIRST PRINTED
PROCESS START               THEN ZBIE STARTS PROCESSING.
TRY PATTERN
P
NOT MATCHED.                (SPEAKING MAN) IS NOT IN A2.
TRY LEARN FROM PATTERN
P
(Z ZDES1 )                  Z STANDS FOR THE UNTRANSLATED
PUT INTO VOCABULARY         (BE(SPEAKING MAN)) FL COMPLEX.
(BE (SPEAKING MAN ))
A1                          THE RUSSIAN LETTER (=I).
P
PUT INTO VOCABULARY         PATTERN P1 IS BEING BUILT.
(SPEAKING MAN )
```

A1 THE RUSSIAN LETTER (=I).

A2 A SET NAME.

P1

NEW PATTERN

<31>	<9-1 = 25418>	<9-2 = 25752>	<9-3 = 252?->
02 9-1	04 J	04 O	04 O
00 A1	00 P0	00 P0	00 Y2
00 Y1	02 9-2		00 Y3
00 A2	00 Y0		
00 Y2	00 J3		
00 A3	00 J4		
00 Y3	00 J3		
	00 T0		
	02 9-3		

COMMENT: PATTERN P1 HAS A P-LIST IDENTICAL TO THE P-LIST OF P0. HOWEVER, THE TRANSLATION RULES ARE DIFFERENT. ZBIE USED (BE(MAN)) → ON, AND (BE (SPEAKING MAN)) → A1, BOTH TRANSLATIONS IN THE CONTEXT OF THE PATTERN P0, FOR MATCH BACK. THE COMPARISON OF THE NON-COMMON PARTS IN FL AND NL GAVE (SPEAKING MAN) → A1 (IN THE CONTEXT A2, P1). SINCE THE NL STRINGS USED FOR MATCH-BACK HAVE NOTHING IN COMMON, 'BE' IS ASSUMED NOT TO BE TRANSLATED, AND THE EXTRACTOR Y1 OF THE SET A1 (WHICH CONTAINS 'BE') IS NOT USED IN THE TRANSLATION RULE T0(P1). IT IS WELL WORTH COMPARING THIS PATTERN P1 TO THE PATTERN P1 CREATED BY ZBIE WHEN LEARNING SOME GERMAN (SEE APPENDIX B).

ZBIE ALSO NOTES THAT IT IS A 'GOOD GUESS' TO ASSUME THAT THE TRANSLATIONS OF A MEMBER OF A3 IN THE CONTEXTS (A3,P0) AND (A3, P1) ARE THE SAME.

```
LOOKING AT SENTENCE 4
(BE (SPOKEN BOY )HERE )
TI2 ZDES1
PROCESS START
TRY PATTERN
P1
NOT MATCHED.
TRY PATTERN
P
NOT MATCHED.
TRY LEARN FROM PATTERN
P1
(Z Z1)                      (SPOKEN BOY) IS NOT IN SET A2,
GUESS                       'HERE' IS NOT KNOWN IN THE CONTEXT
Z1                          (A3,P1), BUT IS GUESSED TO BE
ZDES1                       THE SAME AS IN THE CONTEXT (A3,P0).
PUT INTO VOCABULARY         THE TRANSLATION IS CONSISTENT.
(SPOKEN BOY )
TI2
A2
P1
```

```
PUT INTO VOCABULARY
HERE
ZDES1
A3                          NOW 'HERE' IS ALSO KNOWN IN
P1                          THE CONTEXT (A3,P1).

LOOKING AT SENTENCE 5
(BE (MAN )HERE )
ON ZDES1                    THE FIRST SENTENCE AGAIN.
PROCESS START
TRY PATTERN
P1
PATTERN MATCHED. RESULT:
(Z ZDES1 )                  (MAN) IS NOT KNOWN IN
TRY PATTERN                 THE CONTEXT (A2, P1).
P
PATTERN MATCHED. RESULT:
(ON ZDES1 )
TRY LEARN MORE              PATTERN P1 WAS MATCHED ENTIRELY.
TRY LEARN FROM PATTERN
P1
(Z ZDES1 )
PUT INTO VOCABULARY
(MAN )
ON
A2                          THE SENTENCE WILL BE TRANSLATED
```

```
P1                              BY THE PATTERN P1 FROM NOW ON.

LOOKING AT SENTENCE 6
(BE (SPEAKING MAN )MAN )
A1 MUGYINA
PROCESS START
TRY PATTERN
P1                              'MAN' IS NOT IN THE SET A3.
NOT MATCHED.                    SINCE P0 HAS THE SAME P-LIST AS
TRY PATTERN                     P1, THE MATCH ON P0 IS NOT
P                               ATTEMPTED.
NOT MATCHED.
TRY LEARN FROM PATTERN
P1
(A1 Z )
PUT INTO VOCABULARY
MAN
MUGYINA
A3
P1

LOOKING AT SENTENCE 7
(BE (MAN )MAN )
ON MUGYINA                      A NEW SENTENCE.
PROCESS START
TRY PATTERN
P1
```

```
PATTERN MATCHED. RESULT:
(ON MUGYINA )                    TRANSLATED COMPLETELY.

LOOKING AT SENTENCE 8
(BE (SPOKEN BOY )BOY )
TI2 MAL1YIK
PROCESS START
TRY PATTERN
P1
NOT MATCHED.
TRY PATTERN
P
NOT MATCHED.
TRY LEARN FROM PATTERN
P1
(TI2 Z )
PUT INTO VOCABULARY
BOY                              BUILDING UP THE VOCABULARY.
MAL1YIK
A3
P1

LOOKING AT SENTENCE 9
(BE (SPOKEN GIRL )GIRL )
TI2 DEVOYKA
PROCESS START                    (SPOKEN GIRL) IS NOT IN SET A2,
TRY PATTERN                      'GIRL' IS NOT IN SET A3.
```

```
P1
NOT MATCHED.
TRY PATTERN
P
NOT MATCHED.
TOO HARD
```

TWO MISTAKES AT ONE LEVEL WERE OBTAINED DURING PATTERN MATCHING, SO THAT NO PATTERN IS CLOSE' TO THE FL SENTENCE. THE TWO FL UNITS (SPOKEN GIRL) AND 'GIRL' CANNOT BE TRANSLATED IN ANY CONTEXT (AT THIS STAGE), AND NO NEW PATTERN CAN BE BUILT TO TRANSLATE THE SENTENCE. NOTE THAT IF ZBIE HAD USED SOME OF THE SEMANTICS BUILT INTO FL TO GUESS THAT (SPOKEN GIRL) MAY WELL HAVE A TRANSLATION IDENTICAL TO (SPOKEN BOY), IN THE SAME CONTEXTS, THEN THE SENTENCE COULD HAVE BEEN PROCESSED SUCCESSFULLY AT THIS STAGE. THE GUESSED TRANSLATION WOULD HAVE BEEN 'TI2 7', WHICH IS CONSISTENT WITH THE INPUT.

WE SHALL HEREAFTER LEAVE OJT THE MONOTONOUS 'TRY PATTERN .. PJ ... NOT MATCHED ...'.

```
LOOKING AT SENTENCE 10
(BE BOY )                    THIS SENTENCE MATCHES NO
E2TO MAL1YIK                 PREVIOUS PATTERNS.
PUT INTO VOCABULARY          CREATING PATTERN P2.
BE
E2TO
```

A4 A NEW SET.

P2

PUT INTO VOCABULARY

BOY

MAL1YIK

A3 THE SAME TRANSLATION AS

P2 IN THE CONTEXT (A3,P1).

NEW PATTERN

<P2>	<9-1 = 26178>	<9-2 = 26170>	<9-3 = 26200>
02 9-1	04 0	04 0	04 0
00 A4	00 T0	00 Y1	00 Y1
00 Y1	02 9-2	00 Y2	00 Y2
00 A3	00 D12		
00 Y2	02 9-3		
	00 D13		
	04 26180		

COMMENTS: THE SENTENCE IN FL HAS A LENGTH OF TWO, PREVIOUS PATTERNS EXPECT AN FL SENTENCE OF LENGTH THREE. TO BUILD THE NEW PATTERN, ZBIE TRIED TO FIND TRANSLATIONS, FOR THE FL UNITS IN THE FL SENTENCE, WHICH WERE IDENTICAL OR CLOSE TO NL STRINGS IN THE INPUT NL. 'BE' HAD NO TRANSLATION; HOWEVER, 'BOY' HAD A TRANSLATION 'MAL1YIK', IN THE CONTEXT (A3, P1) (SEE SENTENCE 8), WHICH IS FOUND IN THE INPJT. THE FL UNIT 'BE' WHICH HAD NOT BEEN ACCOUNTED FOR IS MATCHED TO THE REMAINING PARTS OF THE NL INPUT, NAMELY 'E2TO'. ZBIE ALSO MAKES A NOTE THAT IT IS A 'GOOD GUESS'

TO ASSUME THAT THE TRANSLATIONS OF AN FL UNIT IN THE CONTEXTS (A3,P2) AND (A3,P1) ARE IDENTICAL.

```
LOOKING AT SENTENCE 11
(BE FOOT )
E2TO NOG1A                    BUILDING UP THE VOCABULARY.
PROCESS START
TRY LEARN FROM PATTERN
P2
(E2TO Z )
PUT INTO VOCABULARY
FOOT
NOG1A
A3
P2

LOOKING AT SENTENCE 12
(BE FOOT (OF BOY ))           NOT A LINEAR FL SENTENCE.
E2TO NOG1A MAL1YIKA
PROCESS START
TRANSLATE
P2
RESULT:                       Z STANDS FOR THE TRANSLATION OF
(E2TO Z )                     THE FL COMPLEX FOOT(OF BOY), AND,
TRY                           BY THE CONSISTENCY TEST, Z IS TO BE
P2                            REPLACED BY 'NOG1A MAL1YIKA'.
```

```
PUT INTO VOCABULARY          BUILDING SUBPATTERNS P39 AND P38.
BOY
MAL1YIKA
A6                           A DIFFERENT TRANSLATION,
P38                          USING A DIFFERENT SET.
PUT INTO VOCABULARY
FOOT
NOG1A
A3                           THE SAME TRANSLATION AS IN
P39                          THE CONTEXT (A3,P2).
PUT INTO SET
P38                          THE SUBPATTERN P38 IS INSERTED
A7                           AT THE TOP OF THE SET A7.
NEW PATTERN
```

<P38>	<9-1 = 26474>	<9-2 = 26766>	<9-3 = 26660>
02 9-1	04 0	04 0	04 0
00 A5	00 T0	00 Y100	00 Y100
00 Y101	02 9-2		
00 A6	00 D12		
00 Y100	02 9-3		
	00 D13		
	04 26646		
	00 D9		
	00 A7		
	00 V0		
	00 J4		

<P39>	<9-1 = 26454>	<9-2 = 26788>	<9-3 = 24692>
02 9-1	04 0	04 0	04 0
00 A7	00 T0	00 Y98	00 Y98
00 Y99	02 9-2	00 Y99	00 Y99
00 A3	00 D12		
00 Y98	02 9-3		
	00 D13		
	04 26678		
	00 V0		
	00 J4		

PUT INTO SET

P39

A3

COMMENTS: LET US CAREFULLY GO OVER OUR FIRST ENCOUNTER WITH A NON-LINEAR FL SENTENCE. AT THE TOP LEVEL, THE FL SENTENCE MATCHES PARTLY THE PATTERN P2: 'BE' IS AN ELEMENT OF A4, BUT THE SET A3 HAS NO SUBPATTERN MEMBERS. CONSEQUENTLY, THE FL COMPLEX FOOT(OF BOY) CANNOT BE MATCHED.

THE TRANSLATIONS IN PARALLEL OF PATTERN LISTS HAVING MATCHED THE FL SENTENCE TO THE DEEPEST MATCH-DEPTH ARE TRIED. THIS ACTION FOLLOWS 'TRANSLATE'. HERE, ONLY ONE PATTERN, P2, IS FOUND. THEN, THE PATTERN LISTS WITH A CONSISTENT TRANSLATION ARE TRIED. THIS ACTION FOLLOWS 'TRY'. HERE WE ONLY 'TRY' THE PATTERN P2. BY THE CONSISTENCY TEST, THE FL COMPLEX FOOT(OF BOY) CORRESPONDS TO THE NL STRING 'NOG1A MAL1YIKA', AND THE PATTERN CREATING ROUTINE

MANAGES TO BUILD SUBPATTERNS TO MATCH THE FL COMPLEX. NOTE THAT PREVIOUS KNOWLEDGE HAS BEEN USED TO AVOID BUILDING A WHOLE NEW TOP PATTERN.

WHEN BUILDING THE SUBPATTERNS, THE TRANSLATION OF 'FOOT' IN THE CONTEXT (A3,P2) IS USED. THE PREVIOUSLY ENCOUNTERED TRANSLATION OF 'BOY', 'MAL1YIK', IS NOT FOUND IN THE NL STRING HOWEVER, 'MAL1YIK' IS VERY CLOSE TO 'MAL1YIKA', SO THAT THE TRANSLATION OF 'BOY', IN A NEW CONTEXT, IS ASSUMED TO BE 'MAL1YIKA'. SINCE ALL THE WORDS IN THE NL STRING 'NOG1A MAL1YIKA' HAVE BEEN ACCOUNTED FOR, 'OF' IS ASSUMED NOT TO BE TRANSLATED.

THE P-LIST OF THE SUBPATTERN P39 CONTAINS TWO SETS. THE SET A3 CONTAINS FOOT (AND MANY OTHER FL UNITS); THE SET A7 CONTAINS THE SUBPATTERN P38. P38 MATCHES THE FL COMPLEX STRUCTURE (OF BOY), SINCE 'OF' IS A MEMBER OF THE SET A5 AND 'BOY', OF THE SET A6.

LET US VISUALIZE THE SAME SENTENCE BEING MATCHED AGAIN TO THE PATTERN LIST (P2 P39 P38). A MATCH OF THE FL COMPLEX STRUCTURE FOOT(OF BOY) IS ATTEMPTED AGAINST THE SUBPATTERN P39, MEMBER OF A3. A MATCH OF THE FL COMPLEX STRUCTURE (OF BOY) IS ATTEMPTED AGAINST THE SUBPATTERN P38, MEMBER OF A5. ALL MATCHES SUCCEED. THE TRANSLATION RULE OF P2 CALLS FOR THE TRANSLATION OF THE STRUCTURE FOOT(OF BOY) MATCHED AGAINST P39. THE TRANSLATION RULE OF P39 CALLS FOR THE TRANSLATION OF THE ELEMENT MATCHED TO A3, 'FOOT', IN THE CONTEXT (A3,P39), WHICH IS 'NOG1A' (Y98),

FOLLOWED BY THE TRANSLATION OF THE FL COMPLEX (OF BOY) MATCHED TO P38, (Y99). THE TRANSLATION RULE OF P38 CALLS FOR THE TRANSLATION OF THE ELEMENT MATCHED TO A6, 'BOY', IN THE CONTEXT (A6,P38), NAMELY 'MAL1YIKA', (Y100). THE TRANSLATION ROUTINE IS SEEN TO CALL ON ITSELF RECURSIVELY.

NOTE THAT WHILE THE PATTERN P39 IS A MEMBER OF THE SET A3, THE P-LIST OF P39 CONTAINS THE SAME SET A3, SO THAT AN INFINITELY RECURSIVE PATTERN-TREE IS ALREADY POSSIBLE. WHEN USING MORE THAN ONE TOKEN OF THE SAME PATTERN, CARE MUST BE TAKEN TO USE, ALSO, TOKENS OF THE EXTRACTORS.

PATTERN P38, WHICH MATCHES (OF BOY), IS A MEMBER OF THE SET A7. A7 OCCURS ABOVE THE SET A3 (WHICH CONTAINS 'BOY') IN THE P-LIST OF THE PATTERN P39. THIS PECULIARITY IS DUE TO THE IPL-V IMPLEMENTATION OF THE SQUARE-BRACKETED DESCRIPTION LIST IN FL, AS DESCRIBED IN CHAPTER II, PART A.

```
LOOKING AT SENTENCE 13
(BE HAND (OF BOY ))
E2TO RUKA MAL1YIKA
PROCESS START
TRANSLATE
P2                              THE PATTERN LIST.
P39
P38
RESULT:
```

```
(E2TO Z MAL1YIKA )          'HAND' IS NOT IN SET A3.
TRY
P2
P39
P36
PUT INTO VOCABULARY
HAND
RUKA
A3
P39

LOOKING AT SENTENCE 14
(BE HAND )
E2TO RUKA
PROCESS START
TRY PATTERN
P2
PATTERN MATCHED. RESULT:
(E2TO Z )                   'HAND' IS NOT KNOWN IN THE
TRY LEARN FROM PATTERN      CONTEXT (A3,P2).
P2
(E2TO Z )
GUESS                       BUT 'HAND' CAN BE GUESSED, USING
Z                           THE SIMILAR CONTEXTS (A3,P2)
RUKA                        AND (A3,P39).

LOOKING AT SENTENCE 15
```

```
(8E BOOK )
E2TO KNIG1A
PROCESS START
TRY LEARN FROM PATTERN
P2
(E2TO Z )
PUT INTO VOCABULARY
BOOK
KNIG1A
A3
P2                          MORE VOCABULARY.
LOOKING AT SENTENCE 16
(Q BE BOOK WHERE )          Q MEANS QUESTION; IT IS A
G1DE KNIG1A                 MARKER, AND IS NOT TO BE
PROCESS START               TRANSLATED.
PUT INTO VOCABULARY         BUILDING PATTERN P3.
WHERE
G1DE
A10
P3
PUT INTO VOCABULARY
BOOK
KNIG1A
A3
P3
```

<P3>			
02 9-1	04 0	04 0	04 0
00 A8	00 T0	00 Y4	00 Y4
00 Y1	02 9-2	00 Y3	00 Y3
00 A9	00 D12		
00 Y2	02 9-3		
00 A3	00 D13		
00 Y3	04 26952		
00 A10			
00 Y4			

COMMENT: NOTICE THE TRANSLATION RULE.

LOOKING AT SENTENCE 17

(BE (IN (BOOK)HAND))

ONA V RUKE

PROCESS START

TRANSLATE

P2

RESULT:

(E2TO Z) NOT CONSISTENT.

TOO HARD

BOTH (BOOK) AND 'IN' ARE UNKNOWN, AND A NEW TOP PATTERN COULD NOT BE CREATED. ZBIE WILL NOW LEARN (BOOK), THEN RETURN TO THE SAME SENTENCE. 'HAND' COULD BE GUESSED FROM 'RUKA' TO 'RUKE'

```
LOOKING AT SENTENCE 18
(BE (BOOK )WHERE )
ONA ZDES1
PROCESS START
TRY LEARN FROM PATTERN
P1
(Z ZDES1 )
PUT INTO VOCABULARY
(BOOK )
ONA
A2                          NOW (BOOK) IS KNOWN AS 'ONA'
P1                          IN THE CONTEXT (A2,P1).

LOOKING AT SENTENCE 19
(BE (IN (BOOK )HAND ))
ONA V RUKE                  IDENTICAL TO SENTENCE 17.
PROCESS START
TRANSLATE
P2
RESULT:
(E2TO Z )                   NOT CONSISTENT.
PUT INTO VOCABULARY         BUILDING PATTERN P4.
HAND
RUKE
A13                         A NEW SET.
```

P37

PUT INTO VOCABULARY

IN

V

A12

P37

PUT INTO VOCABULARY

(BOOK)

ONA

A2

P37

PUT INTO SET

P37

A14

NEW PATTERN

<P37>

02 9-1

00 A12

00 Y97

00 A2

00 Y96

00 A13

00 Y95

<9-1 = 26506>

04 0

00 T0

02 9-2

00 D12

02 9-3

00 D13

04 27144

00 D9

00 A14

00 V0

<9-2 = 27334>

04 0

00 Y96

00 Y97

00 Y95

<9-3 = 27144>

04 0

00 Y96

00 Y97

00 Y95

00 J4

<P4>	<9-1 = 27164>	<9-2 = 27350>	<9-3 = 27194>
02 9-1	04 0	04 0	04 0
00 A11	00 T0	00 Y2	00 Y2
00 Y1	02 9-2		
00 A14	00 D12		
00 Y2	02 9-3		
	00 D13		
	04 27162		

COMMENT: PATTERNS P2 AND P4 BOTH HAVE TWO SETS ON THEIR P-LISTS, BUT THESE SETS AND THE TRANSLATION RULES OF THE PATTERNS ARE QUITE DIFFERENT. THE TRANSLATION RULE OF P37 IS WORTH NOTING.

LOOKING AT SENTENCE 20
(BE (IN (BOOK)HAND (OF BOY)))
ONA V RUKE MALIYIKA
PROCESS START
TRANSLATE
P4
P37
RESULT:
(ONA V Z)

A CONSISTENT TRANSLATION.
HAND(OF BOY) CORRESPONDS TO
RUKE MALIYIKA.

TRY
P4
P37

PUT INTO VOCABULARY BUILDING SUBPATTERNS P36 AND P35.

HAND

RUKE

A13 NOT A NEW SET(SEE SENTENCE 19).

P36

PUT INTO VOCABULARY

BOY

MALIYIKA

A6 NOT A NEW SET(SEE SENTENCE 12).

P35

PUT INTO SET

P35

A16

NEW PATTERN

<P35>	<9-1 = 27532>	<9-2 = 27776>	<9-3 = 27684>
02 9-1	04 0	04 0	04 0
00 A15	00 T0	00 Y93	00 Y93
00 Y94	02 9-2		
00 A6	00 D12		
00 Y93	02 9-3		
	00 D13		
	04 27674		
	00 D9		
	00 A16		
	00 V0		
	00 J4		

<P36>	<9-1 = 26658>	<9-2 = 27786>	<9-3 = 27720>
02 9-1	04 0	04 0	04 0
00 A16	00 T0	00 Y91	00 Y91
00 Y92	02 9-2	00 Y92	00 Y92
00 A13	00 D12		
00 Y91	02 9-3		
	00 D13		
	04 27706		
	00 V0		
	00 J4		

```
LOOKING AT SENTENCE 21
(BE PENCIL )
E2TO KARANDAW
PROCESS START
TRY LEARN FROM PATTERN
P4                          WHEN THE TRANSLATION IS JUST Z.
(Z )                        WE WAIT THE FIRST TIME AROUND.
WAIT
TRY LEARN FROM PATTERN
P2
(E2TO Z )                   CONSISTENT TRANSLATION.
PUT INTO VOCABULARY
PENCIL
KARANDAW
A3
```

P2

IF WE HAD PROCESSED P4 FIRST, WE WOULD HAVE OBTAINED:

PUT INTO VOCABULARY

PENCIL

E2TO KARANDAW

A14

P4

LOOKING AT SENTENCE 22

(Q BE PENCIL WHERE)

G1DE KARANDAW

PROCESS START

TRY PATTERN

P3

(G1DE Z)

TRY LEARN FROM PATTERN

P3

(G1DE Z)

GUESS

Z

KARANDAW

PUT INTO VOCABULARY

PENCIL

KARANDAW

A3

P3

'PENCIL' IS NOT KNOWN IN THE CONTEXT (A3,P3).

HOWEVER, THE TRANSLATION OF 'PENCIL' CAN BE GUESSED FROM THE CONTEXT (A3,P2) TO THIS CONTEXT, (A3,P3).

IF 'GOOD GUESSES' HAD BEEN USED, THIS SENTENCE WOULD HAVE BEEN TRANSLATED.

```
LOOKING AT SENTENCE 23
(BE (PENCIL )THERE )
ON TAM
PROCESS START
TRY LEARN FROM PATTERN
P1
(Z Z1 )                          THE GUESS MAKES THE TRANSLATION
GUESS                            CONSISTENT WITH THE INPUT.
Z1
TAM
PUT INTO VOCABULARY
(PENCIL )
ON
A2
P1                               MORE VOCABULARY.
PUT INTO VOCABULARY
THERE
TAM
A3
P1

LOOKING AT SENTENCE 24
(BE (IN (PENCIL )DRAWER ))
```

```
ON V A1W2IKE
PROCESS START
TRANSLATE
P4
P37
RESULT:
(Z V Z1 )
GUESS
Z
ON
TRY
P4
P37
PUT INTO VOCABULARY
(PENCIL )
ON
A2
P37
PUT INTO VOCABULARY
DRAWER
A1W2IKE
A13
P37

LOOKING AT SENTENCE 25
(BE BOY (MOD THIS )HERE )
```

EVEN WITHOUT THE GUESS, THE TRANSLATION IS CONSISTENT WITH THE NL INPUT.

E2TOT MAL1YIK ZDES1

PROCESS START

TRANSLATE

P1

RESULT1

(Z ZDES1)

TRY

P1

PUT INTO VOCABULARY BUILDING SUBPATTERNS P34 AND P33.

THIS

E2TOT

A18

P33

PUT INTO VOCABULARY

BOY

MAL1YIK

A3

P34

PUT INTO SET

P33

A19

NEW PATTERN

<P33>	<9-1 = 27360>	<9-2 = 27812>	<9-3 = 2[illegible]>
02 9-1	04 0	04 0	04 0
00 A17	00 T0	00 Y89	00 Y89
00 Y90	02 9-2		

```
00 A18
00 Y89
```

```
00 D12
02 9-3
00 D13
04 26168
00 D9
00 A19
00 V0
00 J4
```

```
<P34>
02 9-1
00 A19
00 Y88
00 A3
00 Y87
```

```
<9-1 = 26422>
04 0
00 T0
02 9-2
00 D12
02 9-3
00 D13
04 27722
00 V0
00 J4
```

```
<9-2 = 27840>
04 0
00 Y88
00 Y87
```

```
<9-3 = 27524>
04 0
00 Y88
00 Y87
```

```
PUT INTO SET
P34
A2

LOOKING AT SENTENCE 26
(BE GIRL )
E2TO DEVOYKA
PROCESS START
```

```
TRY LEARN FROM PATTERN
P4
(Z )
WAIT
TRY LEARN FROM PATTERN
P2
(E2TO Z )
PUT INTO VOCABULARY
GIRL
DEVOYKA
A3
P2                          MORE VOCABULARY.

LOOKING AT SENTENCE 27
(BE (SPOKEN GIRL )GIRL )
TI2 DEVOYKA                 THIS IS SENTENCE 9, WHICH HAD
PROCESS START               BEEN FOUND 'TOO HARD' PREVIOUSLY.
TRY LEARN FROM PATTERN
P1                          NOW, WE USE THE TRANSLATION OF
(Z Z1 )                     'GIRL' IN THE CONTEXT (A3,P2)
GUESS                       AS A 'GUESS' (SEE SENTENCE 26).
Z1
DEVOYKA
PUT INTO VOCABULARY
(SPOKEN GIRL )
TI2
```

```
A2
P1
PUT INTO VOCABULARY
GIRL
DEVOYKA
A3
P1

LOOKING AT SENTENCE 28
(BE GIRL (MOD THIS )HERE )
E2TA DEVOYKA ZDES1
PROCESS START
MATCHED PATTERN-LIST
P1                          PATTERN-LIST MATCHING THE FL
P34                         SENTENCE TO A MATCH-DEPTH OF 2.
P33
RESULT:
(E2TOT Z ZDES1 )            NOT CONSISTENT.
TRANSLATE
P1                          PATTERN LIST MATCHING THE FL
P34                         SENTENCE TO A MATCH-DEPTH OF 1.
RESULT:
(Z Z1 ZDES1 )
GUESS                       WITH THE GUESS, THE TRANSLATION
Z1                          IS CONSISTENT WITH THE NL INPUT.
DEVOYKA
```

```
TRY
P1
P34
PUT INTO VOCABULARY
THIS
E2TA
A21
P32
NEW PATTERN
NOT INSERTED                SEE COMMENTS BELOW.
TRANSLATE                   PATTERN LIST MATCHING THE FL
P1                          SENTENCE TO A MATCH-DEPTH OF 0.
RESULT:
(Z ZDES1 )                  CONSISTENT WITH INPUT. Z TAKES
TRY                         THE PLACE OF GIRL(MOD THIS).
P1
PUT INTO VOCABULARY         BUILDING SUBPATTERNS P30 AND P31.
GIRL
DEVOYKA
A3
P31
PUT INTO VOCABULARY
THIS
E2TA
A21
P30
```

```
PUT INTO SET
P30
A23
NEW PATTERN
```

<P30>	<9-1 = 28520>	<9-2 = 28752>	<9-3 = 28650>
02 9-1	04 0	04 0	04 0
00 A22	00 T0	00 Y83	00 Y83
00 Y84	02 9-2		
00 A21	00 D12		
00 Y83	02 9-3		
	00 D13		
	04 28636		
	00 D9		
	00 A23		
	00 V0		
	00 J4		

<P31>	<9-1 = 28392>	<9-2 = 28762>	<9-3 = 28682>
02 9-1	04 0	04 0	04 0
00 A23	00 T0	00 Y82	00 Y82
00 Y82	02 9-2	00 Y81	00 Y81
00 A3	00 D12		
00 Y81	02 9-3		
	00 D13		
	04 28668		
	00 V0		
	00 J4		

PUT INTO SET

P31

A2

HERE IS OUR FIRST ENCOUNTER WITH ZBIE'S LOOK-AHEAD CAPABILITIES. THE FL SENTENCE IF FIRST MATCHED TO A MATCH-DEPTH OF 2 BY THE PATTERN-LIST (P1 P34 P33) BUT THE TRANSLATION IS NOT CONSISTENT WITH THE NL INPUT, SINCE THE FIRST NL WORDS OF THE TRANSLATION AND INPUT ARE DIFFERENT. NEXT, THE FL SENTENCE IS MATCHED TO A MATCH-DEPTH OF 1 BY THE PATTERN-LIST (P1 P34). THE FL STRUCTURE (MOD THIS) HAS NO CORRESPONDING SUBPATTERN TO WHICH IT IS THEN MATCHED AND CONTRIBUTES THE UNKNOWN Z(=Z0) TO THE TRANSLATION OF THE FL SENTENCE. 'GIRL' IS NOT KNOWN IN THE CONTEXT (A3,P34), AND CONTRIBUTES THE UNKNOWN Z1. SINCE THE TRANSLATION OF 'GIRL' CAN BE CORRECTLY GUESSED, THE TRANSLATION OF THE FL SENTENCE IS CONSISTENT WITH THE INPUT. THE UNKNOWN Z TAKES THE PLACE OF THE FL STRUCTURE (MOD THIS) AND, BY THE CONSISTENCY TEST, IS TO BE REPLACED BY THE FL SENTENCE IS FIRST MATCHED TO A MATCH-DEPTH OF 2 BY THE PATTERN-LIST (P1 P34 P33), BUT THE TRANSLATION IS NOT CONSISTENT WITH THE NL INPUT SINCE THE FIRST NL WORDS OF THE TRANSLATION AND INPUT ARE ISOMORPHIC TO P32. (WE WOULD ONLY REQUIRE THAT P32 BE A HOMOMORPHIC IMAGE OF A SUBPATTERN IN A19). P32 IS 'NOT INSERTED' IN A19 AND THE NEXT PATTERN LIST IS TRIED. SENTENCE 30 WILL PROVIDE ANOTHER EXAMPLE OF THE SAME TECHNIQUES.

```
LOOKING AT SENTENCE 29
(BE TREE )
E2TO DEREVO
PROCESS START
TRY LEARN FROM PATTERN
P4
(Z )
WAIT
TRY LEARN FROM PATTERN
P2
(E2TO Z )
PUT INTO VOCABULARY
TREE
DEREVO
A3                              MORE VOCABULARY BEFORE
P2                              THE NEXT SENTENCE.

LOOKING AT SENTENCE 30
(BE TREE (MOD THIS )HERE )
E2TO DEREVO ZDES1
PROCESS START
MATCHED PATTERN-LIST
PUT INTO VOCABULARY
TREE
DEREVO
A3                              MORE VOCABULARY BEFORE
```

```
P2                          THE NEXT SENTENCE.
LOOKING AT SENTENCE 30
(BE TREE (MOD THIS )HERE )
E2TO DEREVO ZDES1
PROCESS START
MATCHED PATTERN-LIST
P1                          'TREE' IS NOT KNOWN IN THE
P32                         CONTEXT (A3,P31),BUT CAN BE
P30                         GUESSED FROM THE CONTEXT
RESULT:                     (A3,P2).
(E2TA Z ZDES1 )             NOT CONSISTENT.
MATCHED PATTERN-LIST
P1                          'TREE' IS NOT KNOWN IN THE
P34                         CONTEXT (A3,P34),BUT CAN BE
P33                         GUESSED FROM THE CONTEXT
RESULT:                     (A3,P2).
(E2TOT Z1 ZDES1 )           NOT CONSISTENT.
TRANSLATE                   PATTERN-LIST WITH
P1                          MATCH-DEPTH OF 1.
P31
RESULT:
(Z Z1 ZDES1 )               CONSISTENT AFTER GUESS.
GUESS
Z1
DEREVO
```

TRANSLATE

P1

P34

RESULT:

(Z2 Z3 ZDES1)

GUESS

Z3

DEREV0

TRY

P1

P31

PUT INTO VOCABULARY

THIS

E2T0

A25

P29

NEW PATTERN

ALSO A PATTERN-LIST WITH MATCH-DEPTH OF 1, TRANSLATED 'IN PARALLEL' WITH THE ABOVE PATTERN-LIST (P1 P31).

CONSISTENT AFTER GUESS.

NO PREFERENCE BETWEEN THE PATTERN-LISTS, SO 'TRY' THE FIRST ONE CONSIDERED FIRST.

BUILDING SUBPATTERN P29.

<P29>

02 9-1

00 A24

00 Y80

00 A25

00 Y79

<9-1 = 29028>

04 0

00 T0

02 9-2

00 D12

02 9-3

00 D13

04 29142

00 V0

<9-2 = 29154>

04 0

00 Y79

<9-3 = 29154>

04 0

00 Y79

00 J4

NOT INSERTED — P29 IS ISOMORPHIC TO P30.

TRY — THE NEXT PATTERN-LIST.

P1

P34

PUT INTO VOCABULARY — BUILDING SUBPATTERN P28.

THIS

E2TO

A25

P28

NEW PATTERN

<P28>	<9-1 = 29148>	<9-2 = 29258>	<9-3 = 29272>
02 9-1	04 0	04 0	04 0
00 A26	00 T0	00 Y77	00 Y77
00 Y78	02 9-2		
00 A25	00 D12		
00 Y77	02 9-3		
	00 D13		
	04 29122		
	00 V0		
	00 J4		

NOT INSERTED — P28 IS ISOMORPHIC TO P33.

TRANSLATE — PATTERN-LIST WITH

P1 — MATCH-DEPTH OF 0.

RESULT:

(Z ZDES1)

```
TRY
P1
PUT INTO VOCABULARY          BUILDING SUBPATTERNS P27 AND P26.
TREE
DEREVO
A3
P27
PUT INTO VOCABULARY
THIS
E2TO                         'E2TO' WAS ALSO THE TRANSLATION
A25                          OF 'BE' IN THE CONTEXT (A4,P2).
P26
PUT INTO SET
P26
A28
NEW PATTERN
```

<P26>	<9-1 = 29082>	<9-2 = 29476>	<9-3 = 29412>
02 9-1	04 0	04 0	04 0
00 A27	00 T0	00 Y75	00 Y75
00 Y76	02 9-2		
00 A25	00 D12		
00 Y75	02 9-3		
	00 D13		
	04 29398		
	00 D9		
	00 A28		

```
                    00 V0
                    00 J4
<P27>               <9-1 = 29150>      <9-2 = 29526>      <9-3 = 29444>
 02 9-1             04 0               04 0               04 0
 00 A28             00 T0              00 Y74             00 Y74
 00 Y74             02 9-2             00 Y73             00 Y73
 00 A3              00 D12
 00 Y73             02 9-3
                    00 D13
                    04 29430
                    00 V0
                    00 J4
```

PUT INTO SET

P27

A2

LOOKING AT SENTENCE 31

(BE BOY (MOD THAT)THERE)

TOT MAL1YIK TAM

PROCESS START

TRANSLATE	PATTERN-LISTS (P1 P27 P26),
P1	(P1 P31 P30) AND (P1 P34 P33) ALL
P27	MATCH THE FIL SENTENCE TO A
P26	MATCH-DEPTH OF 2, AND ARE
RESULT:	TRANSLATED IN PARALLEL.
(Z Z1 TAM)	CONSISTENT AFTER GUESS.

```
GUESS
Z1
MAL1YIK
TRANSLATE
P1
P31
P30
RESULT:
(Z2 Z3 TAM )                CONSISTENT AFTER GUESS.
GUESS
Z3
MAL1YIK
TRANSLATE
P1
P34
P33
RESULT:
(Z4 MAL1YIK TAM )           CONSISTENT WITH INPUT.
TRY                         THE TRANSLATION OF THE PATTERN-
P1                          LIST (P1 P34 P33) IS CLOSEST TO
P34                         THE NL INPUT, AND (P1 P34 P33)
P33                         IS TRIED FIRST.
PUT INTO VOCABULARY
THAT
TOT
A18
```

P33

IF ZBIE HAD PROCESSED THE PATTERN-LISTS, STARTING WITH THOSE OF DEEPEST MATCH-DEPTH, IN THE ORDER IN WHICH THEY ARE CONSIDERED, THEN ZBIE WOULD HAVE STARTED PROCESSING THE PATTERN-LIST (P1 P27 P26). ZBIE WOULD HAVE CREATED THE FOLLOWING VOCABULARY ENTRIES:

THAT TOT A25 P26

BOY MAL1YIK A3 P27

THE LAST ENTRY WOULD NOW ASSURE THE (INCORRECT) TRANSLATION OF (BE BOY(MOD THIS) HERE) AS 'E2TO MAL1YIK ZDES1'. THE NEXT EXAMPLE IS A SLIGHT VARIATION OF THIS ONE.

```
LOOKING AT SENTENCE 32
(BE GIRL (MOD THAT )THERE )
TA DEVOYKA TAM
PROCESS START
MATCHED PATTERN LIST
P1                          'GIRL' IS NOT KNOWN IN THE
P34                         CONTEXT (A3,P34).
P33
RESULT:
(TOT Z TAM )                NOT CONSISTENT.
TRANSLATE                   TRANSLATE PATTERN-LISTS
P1                          (P1 P27 P28) AND (P1 P31 P30)
P27                         IN PARALLEL.
```

```
P26
RESULT:
(Z Z1 TAM )                    CONSISTENT AFTER GUESS.
GUESS
Z1
DEVOYKA
TRANSLATE
P1
P31
P30
RESULT:
(Z2 DEVOYKA TAM )              CONSISTENT.
TRY                            PATTERN-LIST (P1 P31 P30) GIVES
P1                             A TRANSLATION CLOSEST TO THE
P31                            NL INPUT.
PUT INTO VOCABULARY
THAT
TA
A21
P30

LOOKING AT SENTENCE 33
(BE HAND (OF (SPEAKING MAN ))
E2TO MOA1 RUKA
PROCESS START
TRANSLATE
```

```
P2                          MATCH-DEPTH OF 2.
P39
P38
RESULT:
(E2TO RUKA Z )              NOT CONSISTENT.
TRANSLATE
P2                          MATCH-DEPTH OF 1.
P39
RESULT:
(E2TO RUKA Z )              NOT CONSISTENT.
TRANSLATE                   TRANSLATE IN PARALLEL
P4                          P4 AND P2 (MATCH-DEPTH 0).
RESULT:
(Z )
TRANSLATE
P2
RESULT:
(E2TO Z1 )
TRY                         THE TRANSLATION DUE TO P2 IS
P2                          CLOSEST TO THE NL INPUT.
PUT INTO VOCABULARY         BUILDING SUBPATTERNS P25 AND P24.
(SPEAKING MAN )
MOA1
A30
P24
PUT INTO VOCABULARY
```

HAND

RUKA

A3

P25

PUT INTO SET

P24

A31

NEW PATTERN

<P24>	<9-1 = 29396>	<9-2 = 27938>	<9-3 = 28102>
02 9-1	04 0	04 0	04 0
00 A29	00 T0	00 Y71	00 Y71
00 Y72	02 9-2		
00 A30	00 D12		
00 Y71	02 9-3		
	00 D13		
	04 29418		
	00 D9		
	00 A31		
	00 V0		
	00 J4		

<P25>	<9-1 = 27180>	<9-2 = 29464>	<9-3 = 28614>
02 9-1	04 0	04 0	04 0
00 A31	00 T0	00 Y70	00 Y70
00 Y70	02 9-2	00 Y69	00 Y69
00 A3	00 D12		
00 Y69	02 9-3		

```
                00 D13
                04 28906
                00 V0
                00 J4
PUT INTO SET
P25
A3

LOOKING AT SENTENCE 34
(BE HAND (OF (SPOKEN BOY ))))
E2TO TVOA1 RUKA
PROCESS START
TRANSLATE                   TRANSLATE IN PARALLEL
P2                          PATTERN LISTS (P2 P25 P24) AND
P25                         (P2 P39 P38) WHICH MATCHED
P24                         THE FL SENTENCE TO A DEPTH OF 2.
RESULT:
(E2TO Z RUKA )              CONSISTENT.
TRANSLATE
P2
P39
P38
RESULT:
(E2TO RUKA Z1 )             NOT CONSISTENT.
TRY
P2
```

P25

P24

PUT INTO VOCABULARY

(SPOKEN BOY)

TVOA1

A30

P24

LOOKING AT SENTENCE 35

(BE (ON (BOOK)HAND))

ONA NA RUKE

PROCESS START

TRANSLATE

P4

P37

RESULT:

(ONA Z RUKE)

TRY

P4

P37

PUT INTO VOCABULARY

ON

NA

A12

P37

COMMENT: THIS SLIGHTLY UNNATURAL SENTENCE WAS ADDED TO

AFFORD A SMOOTH TRANSITION TOWARDS THE NEXT SENTENCE WITHOUT BUILDING A NEW SUBPATTERN. SEE THE COMMENTS AT THE END OF SENTENCE 36.

```
LOOKING AT SENTENCE 36
(BE (ON (BOOK )TABLE ))
ONA NA STOLE
PROCESS START
TRANSLATE
P4
P37
RESULT:
(ONA NA Z )                    'TABLE' IS NOT KNOWN IN THE
TRY                            CONTEXT (A13,P37).
P37
PUT INTO VOCABULARY
TABLE
STOLE
A13
P37
```

ASSUME THAT: A) THIS SENTENCE HAD BEEN PRESENTED BEFORE SENTENCE 35; B) 'TABLE' WAS KNOWN IN SOME CONTEXT, FOR INSTANCE AS 'STOL' (NOMINATIVE); AND C) 'ON' WAS NOT KNOWN (WHICH WAS THE CASE HERE). THEN ZBIE WOULD HAVE BUILT A NEW SUBPATTERN BY: A) GUESSING 'TABLE' FROM THE PREVIOUS PRINT-NAME; B) PAIRING (BOOK)

WITH ONE ACCEPTABLE TRANSLATION, 'ONA', OF (BOOK); C) ASSIGNING 'ON' TO THE STILL UNACCOUNTED FOR NL STRING 'NA'. THE NEW SUBPATTERN WOULD HAVE BEEN INSERTED IN THE SET A11, ON THE P-LIST OF P4.

ACTUALLY, IT IS NOT NECESSARY TO BUILD A NEW SUBPATTERN. THE NEW FL SENTENCE FITS INTO THE PREVIOUS PATTERN STRUCTURE IF WE ARE CAREFUL TO GIVE ENOUGH INTERMEDIARY SENTENCES, SUCH AS SENTENCE 35.

ZBIE TENDS TO GENERATE TOO MANY SETS AND PATTERNS. IT WOULD BE INTERESTING TO GIVE ZBIE THE CAPABILITY TO RE-ORGANIZE ITS STRUCTURES, FOR EXAMPLE BY MERGING SOME SETS OR SOME PATTERNS.

```
LOOKING AT SENTENCE 37
(FUTURE TAKE (SPEAKING MAN )BOOK )
A1 VOZ1MU KNIG1U              'FUTURE' IS A MARKER IN FL, AND
PROCESS START                 IS NOT TO BE TRANSLATED.
PUT INTO VOCABULARY           BUILDING PATTERN P5.
BOOK
KNIG1U
A34
P5
PUT INTO VOCABULARY
TAKE
VOZ1MU
A33
```

```
    P5
    PUT INTO VOCABULARY
    (SPEAKING MAN )
    A1                          THE RUSSIAN LETTER.
    A2                          THE SET NAME.
    P5
    NEW PATTERN
<P5>                <9-1 = 28970>       <9-2 = 88212>       <9-3 = 29294>
 02 9-1              04 0                04 0                04 0
 00 A32              00 T0               00 Y3               00 Y3
 00 Y1               02 9-2              00 Y2               00 Y2
 00 A33              00 D12              00 Y4               00 Y4
 00 Y2               02 9-3
 00 A2               00 D13
 00 Y3               04 27858
 00 A34
 00 Y4
   END OF RUN
```

CHAPTER IV.

A CRITICAL LOOK AT ZBIE.

UHR'S PROGRAMS (1964) MAY WELL GIVE AN ILLUMINATING CONTRAST TO SOME OF THE MECHANISMS USED BY ZBIE. UHR SHOWS OUTPUTS FOR TWO PROGRAMS AND DESCRIBES A PLANNED THIRD PROGRAM. ALL HIS PROGRAMS LEARN TO TRANSLATE A NATURAL LANGUAGE, NL1, INTO ANOTHER NATURAL LANGUAGE, NL2.

THE FIRST PROGRAM WAS APPLIED TO TRANSLATE GERMAN INTO ENGLISH. WORDS ARE SEPARATED BY BLANKS. THE PROGRAM DOES NOT USE CONTEXT AND RUNS INTO DIFFICULTIES ON THAT ACCOUNT. FOR EXAMPLE, 'DAS' IN THE GERMAN 'DAS IST' OR 'DAS GLAS' BECOMES RESPECTIVELY 'THIS' OR 'THE'. UHR'S FIRST PROGRAM CANNOT HANDLE SUCH A DIFFICULTY.

THE SECOND PROGRAM, WHICH WAS APPLIED TO TRANSLATE ENGLISH INTO FRENCH, IS MORE POWERFUL IN SEVERAL WAYS. WORDS ARE NOT SEPARATED, AND PART OF THE FUNCTION OF THE PROGRAM IS TO SEPARATE IMPORTANT ELEMENTS IN A CONTINUOUS FLOW. FOR INSTANCE, THE 'S' MARKING A PLURAL CAN BE SEPARATED. IT IS ALSO MORE DIFFICULT TO TRANSLATE A LANGUAGE WITH LITTLE INFLECTION, SUCH AS ENGLISH, INTO A MORE INFLECTED LANGUAGE, SUCH AS FRENCH, THAN TO TRANSLATE

AN INFLECTED LANGUAGE, GERMAN, INTO ENGLISH. IN THE LATTER CASE, AN INCREASED DICTIONARY OFTEN SUFFICES, WHILE IN THE FORMER CASE, CONTEXT IS ESSENTIAL.

TO OBTAIN CONTEXT, CLASSES ARE BUILT ON ELEMENTS OF NL2 CLASS CLUSTERS ARE USED TO RESOLVE AMBIGUITIES AND GENERALIZE PERMUTATIONS. IN THE GIVEN OUTPUT, THOUGH, THE PROGRAM DOES NOT SEEM TO LEARN THE GENERALIZED PERMUTATION: ADJECTIVE + NOUN → NOUN + ADJECTIVE.

THE THIRD (PROJECTED) PROGRAM IS A STRENGTHENING OF PROGRAM 2. IT CAN GENERATE CLASSES OF CLASSES, I.E. ENCOMPASS GROUPS OF WORDS.

ALL THREE PROGRAMS USE STATISTICAL WEIGHTS FOR LEARNING AND UNLEARNING. IT APPEARS THAT SOME USE IS MADE OF THE CORRESPONDING ORDER OF THE LANGUAGES. IN THE GIVEN EXAMPLE 1, FROM: ICH BIN EIN MANN → I AM A MAN, THE INFERENCES ICH → I, BIN → AM, EIN → A, MANN → MAN SEEM TO BE DRAWN; IN EXAMPLE 3, FROM: IF THE BOY → SI LE GARCON AND THE → LE, THE INFERENCES IF → SI, BOY → GARCON SEEM TO BE DRAWN.

LET US APPLY SOME OF THE ABOVE CRITERIA TO ZBIE. INPUTS TO ZBIE ARE WELL SEPARATED WORDS AND ZBIE'S STRING MANIPULATING CAPABILITIES ARE LIMITED TO CHECKING WHETHER THE FIRST FEW CHARACTERS OF TWO WORDS ARE IDENTICAL OR NOT. ZBIE TRIES TO TRANSLATE FROM A STRONGLY NON-INFLECTED LANGUAGE (FL) INTO ANOTHER LANGUAGE, WHICH WILL USUALLY BE INFLECTED. THE CLASSES

USED BY UHR'S PROGRAMS ARE BUILT ON NL2; THE CORRESPONDING SETS ARE BUILT ON FL, EQUIVALENT TO NL1, BY ZBIE. THE CLASSES OF CLASSES DESCRIBED BY UHR FOR THE PROJECTED PROGRAM MAY CORRESPOND TO THE RECURSIVE HIERARCHY (SUB)PATTERN - SET - SUBPATTERN. ZBIE MAKES FEW ASSUMPTIONS ON WORD ORDER, AND HENCE WAS ABLE TO LEARN THE SENTENCES EXHIBITED IN CHAPTER 3 WHEN THE RUSSIAN SENTENCES HAD ALL BEEN PREVIOUSLY INVERTED.

ALL THE PROGRAMS CONSIDERED BUILD STRUCTURES AT RUN-TIME AND USE THESE STRUCTURES TO LEARN FURTHER, WHICH MAY IMPLY BUILDING NEW STRUCTURES.

WE FEEL THAT USING A UNIFORM, STRUCTURED REPRESENTATION, SUCH AS FL, WHICH IS IN SOME SENSE 'UNDERSTOOD', GIVES ZBIE A GREAT ADVANTAGE OVER A FLAT STRING INPUT. IT DESTROYS THE SYMMETRY THAT UHR'S PROGRAMS POSSESS. IT MAY BE THAT STRUCTURES AS POWERFUL AS (OR EVEN MORE POWERFUL THAN) THOSE BUILT IN FL CAN BE CONSTRUCTED EVENTUALLY BY PROGRAMS STARTING FROM (UNSTRUCTURED) STRINGS.

WE HAVE ALREADY ENCOUNTERED VARIOUS DEFICIENCIES IN ZBIE SOME OF THESE COULD BE SOLVED BY MORE PROGRAMMING; OTHERS NECESSITATE A THOROUGH RETHINKING OF THE LEARNING TASK.

AMONG THE FIRST WE SHOULD MENTION:

- THE INITIALIZATION PHASE IS INFLEXIBLE. SENTENCES COULD BE STORED UNTIL TWO SENTENCES MEET THE CONDITIONS TO INITIALIZE THE

FIRST PATTERN.

- THE TRANSLATION RULES SHOULD BE MORE GENERAL. WE GAVE IN CHAPTER II, PART B, TWO KINDS OF DESIRABLE TRANSLATION RULES WHICH ARE PRESENTLY NOT USED BY ZBIE.

- THE METHODS TO OBTAIN 'GOOD GUESSES' SHOULD BE IMPROVED. THESE METHODS COULD USE, FOR INSTANCE, SOME OF THE SEMANTICS BUILT INTO FL (SEE BELOW).

- THE STRING MANIPULATING CAPABILITIES SHOULD BE MORE VERSATILE. IN PARTICULAR, PREFIXES AND SUFFIXES SHOULD BE DISCOVERED, THEREBY ALLOWING GENERALIZATIONS OF TRANSLATIONS. FOR EXAMPLE: THE TRANSLATION OF AN ELEMENT FLI IN THE CONTEXT AI, PI IS OBTAINED BY ADDING SOME STRING IN NL TO THE TRANSLATION OF THE SAME ELEMENT FLI IN SOME OTHER CONTEXT AJ, PJ.

- A PATTERN SHOULD HAVE SEVERAL TRANSLATION RULES TO REFLECT DIFFERENT WAYS IN WHICH THE SAME SITUATION CAN BE EXPRESSED IN NL.

- AS THE NUMBER OF (SUB)PATTERNS INCREASES, HEURISTICS (MAYBE OF A SEMANTIC NATURE) MUST BE USED TO SPEED UP SEARCH DURING MATCHING.

- PRESENTLY, THE CONTEXT OF THE (MAIN) VERB IS THE WHOLE PATTERN, SO THAT, AS THE SAME VERB (IN FL) CHANGES (IN NL), A NEW TOP PATTERN IS CREATED. IN THIS WAY, THE NUMBER OF PATTERNS INCREASES MUCH TOO RAPIDLY.

BESIDES STRUCTURE, WE TRIED TO BUILD SOME SEMANTICS INTO FL. WE CAN SPECIFY THE REFERENT OF A PRONOUN. WE CAN INDEX, IF NECESSARY, SEVERAL PERSONS APPEARING IN A SITUATION TO DIFFERENTIATE AMONG THEM, JUST AS THEY MAY BE DIFFERENTIATED IN A PICTURE. UNTIL NOW, THOJGH, NO USE HAS BEEN MADE OF THE CONTENT OF THE SITUATIONS.

A FIRST STEP, WHICH WAS NEARLY IMPLEMENTED, IS TO USE FINDINGS SUCH AS: THE SUBJECT PRONOUNS (MAN) AND (BOY) HAVE IDENTICAL TRANSLATIONS IN RUSSIAN. WE COULD USE THIS FINDING TO MAKE 'GOOD GUESSES' ON HOW 'MAN' AND 'BOY' WILL BEHAVE IN SIMILAR SITUATIONS. WE COULD ALSO USE THE FUNCTIONAL LANGUAGE DESCRIPTION OF: (SPEAKING MAN(NUMBER 2)), (SPEAKING (AND MAN WOMAN)), ETC... TO OBTAIN THE NOTION OF (SPEAKING PLURAL).

A SECOND STEP WOULD BE TO PROCESS SEVERAL SITUATIONS WHICH ARE DYNAMICALLY RELATED. SEQUENCES SUCH AS: THE BOOK IS ON THE TABLE, I SHALL TAKE THE BOOK IN MY HAND, I AM TAKING THE BOOK, THE BOOK IS IN MY HAND, IT WAS ON THE TABLE, ETC... ARE USED BY I. A. RICHARDS TO TEACH TENSES.

A SIMPLE SCHEME OF SET INCLUSION TESTS ALONE, AS USED BY ZBIE, IS NOT POWERFUL ENOUGH TO LEARN, FOR EXAMPLE, THE RUSSIAN REFLEXIVE POSSESSIVE SVOI1. A MORE POWERFUL SCHEME WHICH LOOKS AT CORRELATIONS AMONG ELEMENTS IN THE FUNCTIONAL LANGUAGE IS NEEDED: IN THE FL SENTENCE

(PUT (PERSON1) (IN HAT(OF (PERSON2)) DRAWER))

ARE PERSON1 AND PERSON2 THE SAME.

THE ABOVE DEFICIENCIES CAN BE OVERCOME BY MAKING BETTER USE OF FL. WE MUST ALSO QUESTION THE USE OF FL.

THE PURPOSE OF FL WAS DESCRIBED IN CHAPTER II. UNDOUBTEDLY, THE 'VISION OF THE WORLD' AS REPRESENTED BY FL IS VERY CLOSE TO AN INDO-EUROPEAN'S 'VISION OF THE WORLD'. IT MAY BE POSSIBLE THAT A SYSTEM COULD DISCARD THE FUNCTIONAL LANGUAGE AND BOOTSTRAP ITSELF TO LOOK AT THE WORLD IN THE LEARNED NATURAL LANGUAGE.

THE RUNNING SYSTEM.

PROGRAMMING ZBIE HAS BEEN A WORTHWHILE EXPERIENCE. THE PROGRAM GENERATED SOME SURPRISES. THE MECHANISMS FOR AVOIDING ERRORS WERE NOT VERY CLEAR AT THE START. THE ADVANTAGE IN TRANSLATING IN PARALLEL PATTERN LISTS OF A GIVEN DEPTH HAD NOT BEEN FORESEEN; IT WAS THOUGHT THAT A SIMPLE HISTORIC MONITOR: LAST PATTERN LIST OBTAINED, FIRST TRIED, WOULD BE ADEQUATE.

AT PRESENT, THE PATTERN BUILDING ROUTINES CAN HANDLE ONLY FL ELEMENTS WHICH HAVE SINGLE NL WORDS AS TRANSLATIONS. (THIS LIMITATION CAN BE OVERCOME BY ADDITIONAL PROGRAMMING.) CONSEQUENTLY, ZBIE WOULD BRANCH TO AN EMPTY EXIT WHEN TRYING TO PROCESS, IN FRENCH, PENCIL → LE CRAYON (TWO NL WORDS). DEALING WITH ARTICLES, MOSTLY UNNECESSARY BUT CUMBERSOME OBJECTS, POSES SOME PROBLEMS. IF PENCIL → LE CRAYON, AND IF WE ALSO KNOW THAT

BED → LE LIT AND WALL → LE MUR, WE COULD CONCLUDE THAT 'LE' IS AN UNNECESSARY OBJECT. THEN IF WE MEET PENCIL --→ UN CRAYON (POSSIBLY), WE CAN ASSUME THAT 'UN' IS ALSO AN UNNECESSARY OBJECT. THE SCHEME COLLAPSES NEXT, THOUGH, WHEN WE MEET PENCIL --→ SON CRAYON (HIS OR HER PENCIL). HERE A TEACHER MAY WELL BE NEEDED.

A 'GOOD' TEACHING SEQUENCE IS PROBABLY VERY IMPORTANT. EVEN WITH A GOOD SEQUENCE, THOUGH, A TIME WILL COME WHEN SOME BAD GENERALIZATIONS HAVE BEEN MADE, AND THESE MUST BE UNLEARNED. ZBIE IS CAREFUL ENOUGH TO AVOID MISTAKES SO THAT, AT THE ELEMENTARY LEVEL CONSIDERED, ERROR RECOVERY IS NOT NEEDED. WE HAVE NO EXCITING SCHEME TO PROPOSE. FOR ERROR RECOVERY, IT APPEARS THAT SUITABLE CHANGES IN SET MEMBERSHIPS MAY BE SUFFICIENT IN SOME CASES, BUT ADDITIONAL WORK AT A MORE ADVANCED LEVEL OF LANGUAGE PROFICIENCY IS IN ORDER.

CHAPTER V.

ENVOI

ZBIE IS A PROGRAM OF MEDIUM LENGTH (ABOUT 5000 LINES OF IPL-V CODE) WHICH TRIES TO SOLVE SOME ASPECTS OF A TASK THAT MAY OR MAY NOT BE VERY DIFFICULT: NATURAL LANGUAGE LEARNING. ZBIE INTERPRETS THE TASK AS LEARNING TO EXPRESS IN A NATURAL LANGUAGE SITUATIONS AS DESCRIBED IN A UNIFORM, STRUCTURED FUNCTIONAL LANGUAGE, WHICH MAY BE CONSIDERED AS GIVING SOME CONTENT TO THE SITUATIONS, ALTHOUGH ACTUALLY VERY LITTLE OF THE POWER OF THE FUNCTIONAL LANGUAGE IS USED.

TO LEARN A LANGUAGE, ZBIE BUILDS ELEMENTARY STRUCTURES AT RUN-TIME: SETS, PATTERNS, SIMPLE TRANSLATION RULES, AN IN-CONTEXT VOCABULARY. IT IS CAPABLE OF SOMEWHAT IMPROVING ITS STRUCTURE TO THE EFFECT THAT SOME PREVIOUSLY LEARNED INSTANCE CAN BE LEARNED IN A BETTER WAY. IT DOES NOT USE STATISTICAL LEARNING SCHEMES. IT HAS SOME ERROR-AVOIDING CAPABILITY, AND HAS POTENTIAL ERROR-RECOVERY POSSIBILITIES WHICH WERE ACTUALLY NOT NEEDED AT THE MODEST LEVEL CONSIDERED HERE. AS COULD BE EXPECTED, ZBIE USES CONTEXT BOTH TO LEARN AND TO AVOID ERRORS.

GIVEN SITUATIONS ARE PARSED IN THE FUNCTIONAL LANGUAGE, NOT IN THE NATURAL LANGUAGE; THE PARSING GIVES A MEASURE OF HOW CLOSE

A NEW SITUATION IS TO CERTAIN CLASSES OF PREVIOUSLY ENCOUNTERED SITUATIONS. ZBIE TRIES TO MINIMIZE ITS LEARNING AT EACH STAGE BY TRYING TO CAPITALIZE ON THE MAXIMUM AMOUNT OF INFORMATION AVAILABLE FROM PREVIOUS SITUATIONS, MEASURED, QUITE SIMPLY, BY THE DEPTH TO WHICH A NEW SITUATION IS PARSED BY A COLLECTION OF PATTERNS. IN FACT, ZBIE ABANDONS THE LEARNING TASK WHEN FACED WITH SITUATIONS THAT IT CONSIDERS TOO DIFFICULT TO HANDLE AT A GIVEN STAGE. SINCE IT CAN DIAGNOSE THE PARTICULAR DIFFICULTIES ENCOUNTERED, ZBIE COULD ASK APPROPRIATE QUESTIONS IN A TIME-SHARING ENVIRONMENT.

IT WOULD APPEAR THAT THE FUNCTIONAL LANGUAGE SHOULD POSSESS IN ITS DESCRIPTION ALL THE SEMANTIC SUBTLETIES OF THE NATURAL LANGUAGE LEARNT, AN UNAPPEALING FACT. A MUCH MORE POWERFUL SYSTEM MAY BE ABLE TO BOOTSTRAP ITSELF AND USE THE NATURAL LANGUAGE IT HAS STARTED TO LEARN AS ITS MAIN REPRESENTATION, WITH POSSIBLE REFERENCES FROM TIME TO TIME TO THE FUNCTIONAL LANGUAGE. SEMANTIC SUBTLETIES COULD THEN BE DESCRIBED IN THE NATURAL LANGUAGE ITSELF.

IT APPEARS TO BE FRUITFUL TO LOOK AT NATURAL LANGUAGES AS WAYS TO EXPRESS SITUATIONS WITH STRUCTURE AND CONTENT. MUCH VALUABLE RESEARCH CAN BE DONE IN AREAS TOUCHED UPON BY ZBIE, AMONG WHICH THE DESIGN OF MORE POWERFUL, SEMANTICALLY ORIENTED, FUNCTIONAL LANGUAGES; THE PROBLEM OF ERROR RECOVERY, AND THE EVOLUTIONARY REORGANIZATION AND IMPROVEMENT OF STRUCTURES MAY WELL BE THE MOST INTERESTING AND CHALLENGING.

APPENDIX A.

CODE FOR CYRILLIC ALPHABET.

А	A	Р	R
Б	B	С	S
В	V	Т	T
Г	G1	У	U
Д	D	Ф	F
Е	E	Х	H
Е	ET	Ц	C
Ж	G	Ч	Y
З	Z	Ш	W
И	I	Щ	W2
Й	IT	ъ	2
К	K	Ы	I2
Л	L	Ь	T
М	M	Э	E2
Н	N	Ю	UT
О	O	Я	AT
П	P		

APPENDIX B.

'EVOLUTIONARY LEARNING'.

WE ARE PRESENTING HERE A SHORT EXAMPLE, IN GERMAN, EXHIBITING AN ASPECT OF ZBIE'S CAPABILITIES. THE TRANSLATION RULE OF PATTERN P1 IS INCORRECT. AS MORE EXAMPLES ARE GIVEN, THOUGH, ZBIE BUILDS NEW PATTERNS P2 AND P3. THE SENTENCES WHICH WERE PREVIOUSLY TRANSLATED WHEN REACHING PATTERN P1 ARE NOW TRANSLATED WHEN REACHING PATTERNS P2 OR P3, SO THAT PATTERN P1 IS NOT REACHED ANY MORE. PATTERN P1 HAS EFFECTIVELY BEEN WASHED OUT.

THE OUTPUT IS REPRODUCED AS OBTAINED FROM THE PRINTER. INPUTS IN IPL-V FORM HAVE BEEN DELETED EXCEPT IN ONE ILLUSTRATIVE CASE.

THE IMPORTANT ATTRIBJTES AND VALUES OF A PATTERN PJ HAVE THE FOLLOWING MEANING:

TO(PJ) = TRANSLATION RULE (A LIST STRUCTURE).
PO(PJ) = A LIST OF PATTERNS WITH A P-LIST IDENTICAL TO PJ.
THE OTHER ATTRIBUTES ARE USED FOR BOOK-KEEPING AND MAY BE DISREGARDED.

COMMENTS ARE GIVEN AT THE RIGHT OF INPUTS, OR AT THE END OF

THE PROCESSING OF A SENTENCE.

LOOKING AT SENTENCE	COMMENTS
(BE (WOMAN)HERE)	
SIE IST HIER	
(BE (WOMAN)THERE)	
SIE IST DORT	THE FIRST TWO SENTENCES
PROCESS START	TO START THE INITIALIZATION
PUT INTO VOCABULARY	BE IS IN SET A1;(WOMAN) IN SET A2
(BE (WOMAN))	
(SIE IST)	
P	CONTEXT VOCABULARY FOR FL COMPLEX
PUT INTO VOCABULARY	
HERE	IN-CONTEXT VOCABULARY FOR FL
HIER	UNIT. THE FL UNIT (IN THIS CASE,
A3	HERE) IS INSERTED IN THE SET (IN
P	THIS CASE, A3). P=P0.
PUT INTO VOCABULARY	
THERE	
DORT	
A3	
P	
NEW PATTERN	

<P0>	<9-1 = 25256>	<9-2 = 25482>	<9-3 = 255. >
02 9-1	04 0	04 0	04 0
00 A1	00 T0	02 9-3	00 Y1

```
00 Y1          02 9-2          00 Y3          00 Y2
00 A2
00 Y2
00 A3
00 Y3
```

COMMENT: THE INITIALIZATION PHASE IS OVER. FOR THE NEXT SENTENCE, WE SHOW THE IPL-V INPUT.

```
IP INPUT DATA                  5       01
                               X4      95      0
                               95      0
IP BE                                  E0
IP (MAN(NUMB 2))                       91
IP   THERE                             E4      0
IP                             9-1     0
IP                                     0
IP (NUMB 2)                            92
IP MAN                                 E2      0
IP                             92      0
IP NUMB                                E10
IP 2                                   E13     0
IP                             X3      95      0
IP                             95      0
IP SIE                                 R0
IP SIND                                R4
```

```
IP DORT                                          R3     0

IP                                        5      J0

H2 = 2339  CELLS          00:00:06

LOOKING AT SENTENCE
(BE (MAN (NUMB 2 ))THERE )
SIE SIND DORT                    THE INPUTS ARE FIRST PRINTED
PROCESS START
TRY PATTERN
P
NOT MATCHED.
TRY LEARN FROM PATTERN
P                                (BE(MAN(NUMB 2)) IS THE UNTRANS-
(Z DORT)                         LATED PART, CORRESPONDING TO Z
PUT INTO VOCABULARY
(BE (MAN (NUMB 2 )))
(SIE SIND )
P
PUT INTO VOCABULARY              PATTERN P1 IS BEING BUILT
(MAN (NUMB 2 ))
SIND
A2
P1
PUT INTO VOCABULARY
```

BE

SIE

A1

P1

NEW PATTERN

<P1>	<9-1 = 25256>	<9-2 = 25774>	<9-3 = 25422>
02 9-1	04 0	04 0	04 0
00 A1	00 P0	00 P0	00 Y1
00 Y1	02 9-2		00 Y2
00 A2	00 Y0		00 Y3
00 Y2	00 J3		
00 A3	00 J4		
00 Y3	00 J3		
	00 T0		
	02 9-3		

COMMENT: PATTERN P1 HAS A P-LIST IDENTICAL TO THE P-LIST OF P0. HOWEVER, THE TRANSLATION RULES ARE DIFFERENT. ZBIE USED (BE(WOMAN)) → SIE IST, AND (BE(MANINUMB 2)) → SIE SIND, BOTH TRANSLATIONS IN THE CONTEXT OF THE PATTERN P0, FOR MATCH BACK THE COMMON PARTS IN FL AND NL, RESPECTIVELY 'BE' AND 'SIE', WERE PAIRED. THE COMPARISON OF THE NON-COMMON PARTS GAVE (MANINUMB 2) → SIND (CONTEXT A2,P1). THE INFERENCE IS GRAMMATICALLY INCORRECT. NOTE THAT THE TRANSLATION RULE IS (Y1 Y2 Y3), IN THE SAME ORDER AS THE SETS. IT IS WELL WORTH COMPARING THIS PATTERN P1 TO THE PATTERN P1 CREATED BY ZBIE WHEN LEARNING RUSSIAN (SEE CHAPTER

III).

```
LOOKING AT SENTENCE
(BE (WOMAN )HERE )
SIE IST HIER                    THE FIRST SENTENCE AGAIN
PROCESS START
TRY PATTERN
P1
PATTERN MATCHED. RESULT:
(SIE Z Z1 )                     (WOMAN) IS NOT KNOWN IN THE
TRY PATTERN                     CONTEXT (A2,P1), HENCE THE Z
P                               ID. FOR HERE, (A3,P1) AND Z1
PATTERN MATCHED. RESULT:
(SIE IST HIER )
TRY LEARN MORE                  THE PATTERN P1 WAS THE FIRST TO
TRY LEARN FROM PATTERN          BE TOTALLY MATCHED, BUT WAS NOT
P1                              SUCCESSFULLY TRANSLATED
(SIE Z Z1 )
GUESS                           Z (I.E. (WOMAN)) IS NOT GUESSED
Z1                              BUT Z1 IS, GIVING A CONSISTENT
HIER                            TRANSLATION WITH THE INPUT
PUT INTO VOCABULARY
(WOMAN )
IST
A2
P1
```

```
PUT INTO VOCABULARY
HERE                          NOW 'HERE' IS ALSO KNOWN
HIER                          IN THE CONTEXT (A3,P1)
A3
P1
```

COMMENT: NOW (BE (WOMAN) HERE) WILL BE TRANSLATED BY PATTERN P_1.

```
LOOKING AT SENTENCE
(BE (MAN )HERE )
ER IST HIER                   A NEW SENTENCE
PROCESS START
TRY PATTERN
P1
NOT MATCHED.                  (MAN) IS NOT IN A2
TRY PATTERN                   AS THE P-LISTS OF P0 AND P1 ARE
P                             IDENTICAL, IT IS NOT NECESSARY
NOT MATCHED.                  TO ACTUALLY MATCH P0
TRY LEARN FROM PATTERN
P1
(SIE Z HIER)                  NOT CONSISTENT WITH INPUT
TRY LEARN FROM PATTERN
P
(Z HIER)                      Z STANDS FOR THE UNKNOWN
PUT INTO VOCABULARY           TRANSLATION OF (BE(MAN))
(BE (MAN ))
```

(ER IST)

P

PUT INTO VOCABULARY CREATING PATTERN P2

(MAN)

ER

A2

P2

PUT INTO VOCABULARY

BE

IST

A1

P2

NEW PATTERN

<P2>	<9-1 = 25814>	<9-2 = 26054>	<9-3 = 25854>
02 9-1	04 0	04 0	04 0
00 A1	00 P0	00 P0	00 Y2
00 Y1	02 9-2	00 P1	00 Y1
00 A2	00 Y0		00 Y3
00 Y2	00 J3		
00 A3	00 J4		
00 Y3	00 J3		
	00 T0		
	02 9-3		

COMMENT: WE MATCHED BACK (BE(MAN)) → ER IST AND (BE(WOMAN)) → SIE IST, TO OBTAIN A NEW PATTERN P2. THE TRANSLATION RULE OF

P2, (Y2 Y1 Y3), IS DIFFERENT FROM THE TRANSLATION RULE OF P1 AND IS GRAMMATICALLY CORRECT. NOTE THAT THE TRANSLATION RULE CONTAINS A TRANSFORMATION.

```
LOOKING AT SENTENCE
(BE (WOMAN )HERE )
SIE IST HIER                    AGAIN THE FIRST SENTENCE
PROCESS START
TRY PATTERN
P2
PATTERN MATCHED. RESULT:
(Z IST Z1 )
TRY PATTERN
P1
PATTERN MATCHED. RESULT:
(SIE IST HIER)                  NOW TRANSLATED BY P1
TRY LEARN MORE
TRY LEARN FROM PATTERN
P2
(Z IST Z1 )
GUESS
Z1
HIER
PUT INTO VOCABULARY
HERE
HIER
```

```
A3
P2
PUT INTO VOCABULARY
(WOMAN )
SIE
A2                              NOW THE SENTENCE WILL BE
P2                              TRANSLATED BY P2

LOOKING AT SENTENCE
(BE (SPEAKING MAN (NUMB 2 ))HERE )
WIR SIND HIER                   A NEW SENTENCE
PROCESS START
TRY PATTERN
P3                              (SPEAKING MAN(NUMB 2) IS
NOT MATCHED.                    NOT IN A2
TRY PATTERN
P1
NOT MATCHED.
TRY PATTERN
P
NOT MATCHED.
TRY LEARN FROM PATTERN
P2
(Z IST HIER)                    NOT CONSISTENT WITH INPUT
TRY LEARN FROM PATTERN
P1
```

```
(SIE Z HIER )
TRY LEARN FROM PATTERN
P
(Z HIER )
PUT INTO VOCABULARY
(BE (SPEAKING MAN (NUMB 2 )))
(WIR SIND )
P
PUT INTO VOCABULARY          CREATING PATTERN P3
(SPEAKING MAN (NUMB 2 ))
WIR
A2
P3
PUT INTO VOCABULARY
BE
SIND
A1
P3
NEW PATTERN
```

<P3>	<9-1 = 25584>	<9-2 = 26328>	<9-3 = 2613 >
02 9-1	04 0	04 0	04 0
00 A1	00 P0	00 P0	00 Y2
00 Y1	02 9-2	00 P2	00 Y1
00 A2	00 Y0	00 P1	00 Y3
00 Y2	00 J3		
00 A3	00 J4		

```
00 Y3                00 J3
                     00 T0
                     02 9-3
```

COMMENT: THE TRANSLATION RULES OF P3 AND P2 ARE IDENTICAL, BUT 'BE' IS TRANSLATED DIFFERENTLY IN THE CONTEXTS (A1,P2) AND (A1,P3).

```
LOOKING AT SENTENCE
(BE (MAN (NUMB 2 ))THERE )
SIE SIND DORT                    THE THIRD SENTENCE AGAIN
PROCESS START
TRY PATTERN
P3
PATTERN MATCHED. RESULT:
(Z SIND Z1 )
TRY PATTERN
P2
PATTERN MATCHED. RESULT:
(Z IST Z1 )
TRY PATTERN
P1
PATTERN MATCHED. RESULT:
(SIE SIND Z )
TRY PATTERN
P
PATTERN MATCHED. RESULT:
```

```
(SIE SIND DORT )
TRY LEARN MORE
TRY LEARN FROM PATTERN
P3
(Z SIND Z1 )
GUESS                          THE TRANSLATION IS CONSISTENT WITH
Z1                             THE INPUT EVEN WITHOUT THE GUESS
DORT                           OF Z1 (STANDING FOR 'THERE')
PUT INTO VOCABULARY
THERE
DORT
A3
P3
PUT INTO VOCABULARY
(MAN (NUMB 2 ))
SIE
A2                             NOW THE SENTENCE WILL BE
P3                             TRANSLATED BY P3
```

COMMENT: WE SHALL NOW GIVE SENTENCES ONE AND THREE AGAIN, AND SHOW HOW THEY ARE TRANSLATED BY P2 AND P3 RESPECTIVELY. THE PATTERN P1 IS NOT REACHED ANY MORE.

```
LOOKING AT SENTENCE
(BE (WOMAN )HERE )
SIE IST HIER
PROCESS START
```

```
TRY PATTERN
P3
PATTERN MATCHED. RESULT:
(Z SIND Z1 )
TRY PATTERN
P2
PATTERN MATCHED. RESULT:
(SIE IST HIER )
TRY LEARN MORE
TRY LEARN FROM PATTERN
P3
(Z SIND Z1)                    NOT CONSISTENT

LOOKING AT SENTENCE
(BE (MAN (NUMB 2 ))THERE )
SIE SIND DORT
PROCESS START
TRY PATTERN
P3
PATTERN MATCHED. RESULT:
(SIE SIND DORT )
END OF RUN
```

BIBLIOGRAPHY.

CHOMSKY, N. AND MILLER, G. A., 'INTRODUCTION TO THE FORMAL ANALYSIS OF NATURAL LANGUAGES,' IN LUCE, R. D., BUSH, R., AND GALANTER, E., HANDBOOK OF MATHEMATICAL PSYCHOLOGY, VOL. 2, WILEY, NEW YORK, 1963, P. 277.

MCCARTHY, J. ET AL., LISP 1.5 PROGRAMMERS MANUAL, MIT PRESS, CAMBRIDGE, MASS., 1963.

NEWELL, A. ET AL., INFORMATION PROCESSING LANGUAGE-V MANUAL, PRENTICE-HALL, ENGLEWOOD CLIFFS, N.J., 1964.

PIVAR, M. AND FINKELSTEIN, M. 'AUTOMATION, USING LISP, OF INDUCTIVE INFERENCE ON SEQUENCES,' IN THE PROGRAMMING LANGUAGE LISP, INFORMATION INTERNATIONAL, INC., CAMBRIDGE, MASS., 1964, PP. 125-136.

REICHENBACH, H., ELEMENTS OF SYMBOLIC LOGIC, MACMILLAN, NEW YORK. 1947.

RICHARDS, I. A., ET AL., RUSSIAN THROUGH PICTURES, BOOK I, WASHINGTON SQUARE PRESS, NEW YORK, 1961. THERE ARE SIMILAR BOOKS FOR: ENGLISH, FRENCH, HEBREW, ITALIAN, SPANISH AND GERMAN.

SIMON, H. A. AND KOTOVSKY, K., 'HUMAN ACQUISITION OF CONCEPTS FOR

SEQUENTIAL PATTERNS,' PSYCHOLOGICAL REVIEW, VOL. 70, NO 6, 1963, PP. 534-546.

SOLOMONOFF, R. J. 'THE MECHANIZATION OF LINGUISTIC LEARNING,' IN PROCEEDINGS OF THE SECOND INTERNATIONAL CONGRESS ON CYBERNETICS. NAMUR, BELGIUM, 1960, PP. 180-193.

SOLOMONOFF, R. J. ' A FORMAL THEORY OF INDUCTIVE INFERENCE,' INFORMATION AND CONTROL, 1964, PP. 1-22, 224-254.

SOLOMONOFF, R. J. 'SOME RECENT WORK IN ARTIFICIAL INTELLIGENCE,' PROCEEDINGS OF THE IEEE, VOL. 54, NO. 12, DEC. 1966, PP 1687-1697.

UHR, L. 'PATTERN STRING LEARNING PROGRAMS,' BEHAVIORAL SCIENCE, VOL. 9, JULY 1964, PP. 258-270.

Security Classification

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Carnegie-Mellon University Department of Computer Science Pittsburgh, Pennsylvania 15213	UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

NATURAL LANGUAGE LEARNING BY COMPUTER

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

Laurent Siklossy

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
May 1968	109	11

8a. CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO SD-146 9718 c. 6154501R d. 681304	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) OSR 68-1594

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH, OTHER	Air Force Office of Scientific Research (SRIA) 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

Learning a natural language is taken as an improvement in a system's ability to express situations in a natural language.

This dissertation describes a computer program, called Zbie, written in IPL-V, which accepts the description of situations in a uniform, structured functional languageand tries to express these situations in a natural language. Examples are given for German and, mostly, Russian.

At run-time, Zbie builds simple memory structures. Patterns and sets are built on the functional language. The translation rules of the patterns and an in-context vocabulary provide the transition to the natural language. Zbie is a cautious learner, and avoids errors by several mechanisms. Zbie is capable of some evolutionary learning.

DD FORM 1 NOV 65 1473

Security Classification